EXHIBIT B

~~DAVID H. HARPER*~~
~~*david.harper@haynesboon.com*~~
~~JASON P. BLOOM*~~
~~*jason.bloom@haynesboone.com*~~
~~**HAYNES AND BOONE, LLP**~~
~~2801 N. Harwood~~JOSHUA D. BRANSON*
jbranson@kellogghansen.com
DANIEL V. DORRIS*
ddorris@kellogghansen.com
BETHAN R. JONES*
bjones@kellogghansen.com
MATTHEW D. READE*
mreade@kellogghansen.com
TIBERIUS T. DAVIS*
tdavis@kellogghansen.com
KELLOGG, HANSEN, TODD,
FIGEL & FREDERICK, P.L.L.C.
1615 M Street
, N.W., Suite ~~2300~~400
~~Dallas, Texas 75201~~
Washington, D.C. 20036
Telephone: (~~214) 651-5000~~202) 326-7900
Facsimile: (~~214) 651-5940~~202) 326-7999
* *Admitted Pro Hac Vice*

~~JASON T. LAO, SBN 288161~~
~~*jason.lao@haynesboone.com*~~
~~ANDREA LEVENSON, SBN 323926~~
~~*andrea.levenson@haynesboone.com*~~
~~**HAYNES AND BOONE,**~~ ADRIAN SAWYER, State Bar No. 203712
sawyer@sawyerlabar.com
SAWYER & LABAR LLP
~~600 Anton Boulevard~~1700 Montgomery Street, Suite ~~700~~108
~~Costa Mesa~~San Francisco, California ~~92626~~94111
Telephone: (~~949) 202-3000~~415) 262-3820
~~Facsimile: (949) 202-3001~~

Counsel
~~Attorneys~~ for Plaintiff
X Corp.

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

X CORP., a Nevada corporation,

~~Plaintiff,~~

~~vs~~ Plaintiff,

v.

BRIGHT DATA LTD., an Israeli corporation,

Defendant.

Case No. 3:23-cv-03698-WHA

~~FIRST~~[PROPOSED] SECOND AMENDED COMPLAINT

JURY TRIAL ~~DEMAND~~DEMANDED

~~Plaintiff X Corp. ("X Corp." or "Plaintiff"), by and through its undersigned counsel, hereby files its First Amended Complaint against Defendant Bright Data Ltd., ("Bright Data" or "Defendant"), and in support thereof alleges as follows:~~

INTRODUCTION

1. Defendant Bright Data Ltd. ("Bright Data") has built an illicit data-scraping business on the backs of innovative technology companies like X Corp., which operates the social media platform formerly known as Twitter and now known as X. Bright Data scrapes and sells millions of records from ~~X Corp.'s~~the X platform~~, in blatant violation of X Corp.'s Terms of Service, by which Bright Data is bound.~~. Bright Data also ~~induces and facilitates other X users to violate their own agreements with X Corp. by selling automated data~~ sells tools enabling others to unlawfully scrape data from X without (in its own words) being "flagged or blocked" or "detect[ed]." To facilitate the scraping ~~tools~~ it knows X is trying to stop, Bright Data sells tools enabling its customers to "[a]void IP blocks and bans," to "[b]ypass geo-restrictions and ~~services~~CAPTCHAs," and to "overcome anti-bot detection" measures that ~~specifically target a wide range of X Corp. data.~~X employs to safeguard its platform.

2. Bright Data ~~uses elaborate technical~~purports to be the world's largest provider of illicit data scraping. It markets itself as the "leading proxy provider" allowing users around the globe to circumvent technological measures ~~to evade~~that platforms like X ~~Corp.'s anti-scraping technology, taxing the resources of X Corp.'s~~deploy to prevent scraping. It is a substantial source of all data scraping that takes place on X's platform. And it drives massive amounts of scraping activity even as it goes to great lengths to conceal its involvement and prevent X from tracking its activities.

3. X implements stringent technological measures to stop scraping by Bright Data, Bright Data's customers, and others. Those measures restrict much of the content on X's platform to logged-in users, shielding that content from the public writ large. Data scrapers can only access the content they want through logged-in accounts, so scrapers use fake, automatically created accounts to obtain it. Even if Bright Data does not itself scrape such non-public data, it knowingly enables its customers to do so and markets circumvention tools to customers for that purpose.

4. Unlawful data scraping requires X to spend [redacted] buying excess server capacity to absorb the unwanted scraping requests with which Bright Data bombards its servers~~ and hampering~~. Bright Data's circumvention further degrades the user experience on X's platform by spreading fake accounts and promoting spam. It threatens X user privacy by allowing real-time automated surveillance of user activity. And it deprives X of the revenue it would obtain if developers paid for legitimate ~~X users. Bright Data~~access through X's Application Programming Interfaces ("APIs").

5. Bright Data's business depends on circumventing the technological measures X employs to protect its platform and users. That circumvention violates X Corp.'s Terms of Service, by which Bright Data is bound, and induces Bright Data's customers to breach their own similar agreements with X Corp. Bright Data even proudly advertises that it "avoids" and "bypasses" the technological measures X has deployed to protect and control access to data on the X platform – in direct contravention of the Digital Millennium Copyright Act ("DMCA"), Computer Fraud and Abuse Act ("CFAA"), and California's Computer Data Access and Fraud Act ("CDAFA").

~~2.~~6. Bright Data is a sophisticated actor that is aware ~~that~~ its activities violate X Corp.'s Terms of Service, Privacy Policy, and the Rules and Policies (together, the "Terms"), because the company and its executives are (or were) registered X account holders who ~~have~~ agreed to abide by those Terms~~.~~ and thus have full knowledge of them.

~~3.~~7. X Corp. brings this action for injunctive relief to halt Bright Data's unauthorized use of X Corp.'s platform and other unlawful conduct and for damages caused by Bright Data's breach.

THE PARTIES

~~4.~~8. Plaintiff X Corp. is a privately held corporation duly organized and existing under the laws of the State of Nevada with its principal place of business at 1355 Market Street, Suite 900, San Francisco, California, 94103. X Corp. owns and operates the social media platform X, formerly known as Twitter.

~~5.~~9. On information and belief, Defendant Bright Data was incorporated in Israel in 2008 as Zon Networks Ltd. and changed its name to Bright Data Ltd. in 2021. Bright Data has its principal place of business at 4 Hamahshev St., Netanya 4250714, in Israel. Bright Data has at times maintained an office at L415 Mission Street, 37th Floor, in San Francisco, California.

~~6.~~10. Defendant Bright Data operates brightdata.com, where it sells data scraped from numerous websites and social media platforms, including X, along with tools and services to scrape data from X and other platforms.

JURISDICTION AND VENUE

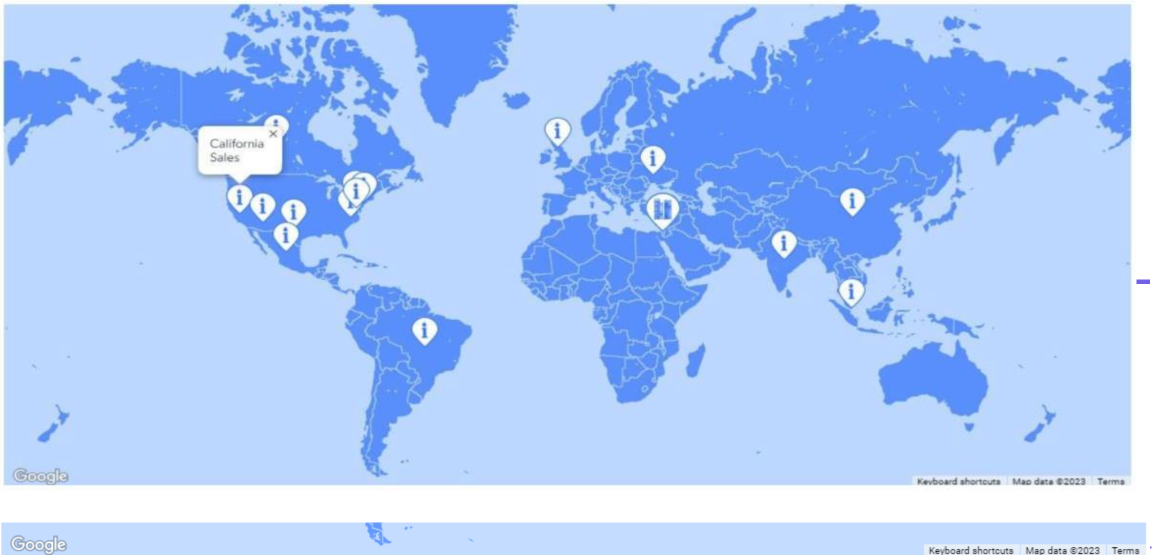
~~7.~~11. This Court has jurisdiction over this action under 28 U.S.C. § 1332 because complete diversity exists, and the amount in controversy exceeds $75,000. Plaintiff X Corp. is incorporated in Nevada with its principal place of business in California. Defendant Bright Data is incorporated in Israel with its principal place of business in Israel.

~~8.~~12. This Court has personal jurisdiction over Defendant because Defendant has consented to X Corp.'s Terms, which require all disputes related to the Terms be brought in the federal or state courts located in San Francisco, California. As part of its agreement to those Terms, Defendant also consented to personal jurisdiction in California.

~~9.~~13. Additionally, this Court has personal jurisdiction over Defendant because Defendant knowingly directed prohibited conduct to California and California residents. Defendant offers its data sets and scraping tools for sale in California and to California residents, and has targeted X Corp., which has its principal place of business in California, as well as X Corp.'s users located in California.

~~10.~~14. Defendant markets and sells its products to California residents and businesses via a sales office in California, according to its website:

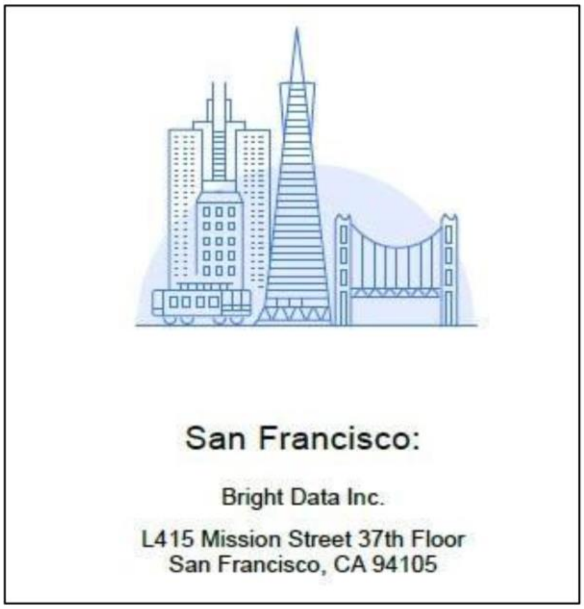
Figure 1: Screenshot of Bright Data's website on November 14, 2023


California
Sales
Google
Keyboard shortcuts | Map data ©2023 | Terms
Google
Keyboard shortcuts | Map data ©2023 | Terms

~~*See* Exh. A.~~

~~11.~~15. As recently as October 19, 2022, Defendant encouraged customers to contact Bright Data at its California sales office, as shown in Figure 2.

Figure 2: Screenshot from Bright Data's "Contact Us" page on October 19, 2022



San Francisco:

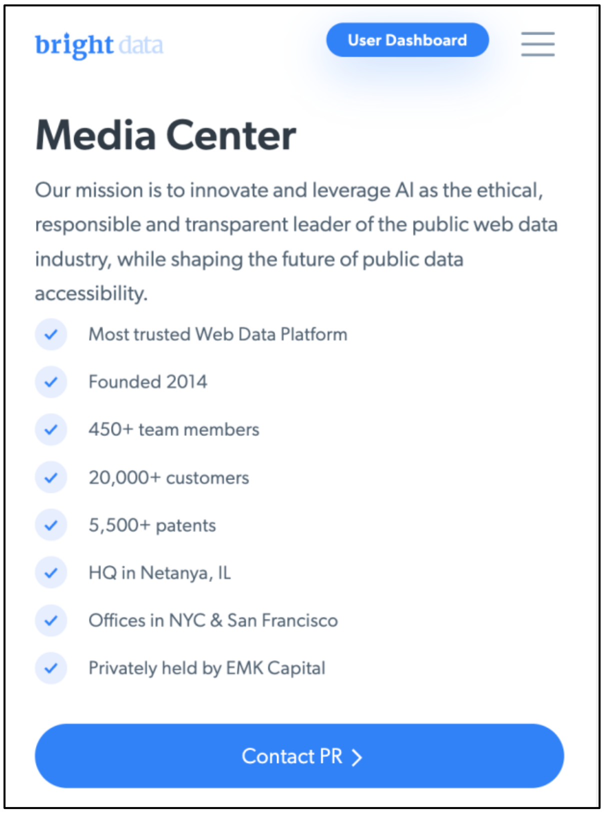
Bright Data Inc.

L415 Mission Street 37th Floor
San Francisco, CA 94105

~~*See* Exh. B.~~

~~12.~~16. Members of Defendant's business development and sales team are also located in California._ For example, Defendant's Chief Revenue Officer, who oversees Bright Data's sales operations, is based in the San Francisco Bay Area. ~~*See* Exh. C.~~ Defendant's Global Head of Presales is also based in California, along with numerous other Bright Data employees. ~~*See* Exh. C~~ To this day, Defendant advertises its San Francisco presence to clients, as shown in Figure 3.

Figure 3: Screenshot from Bright Data's "Media Center" page on August 11, 2024



Media Center

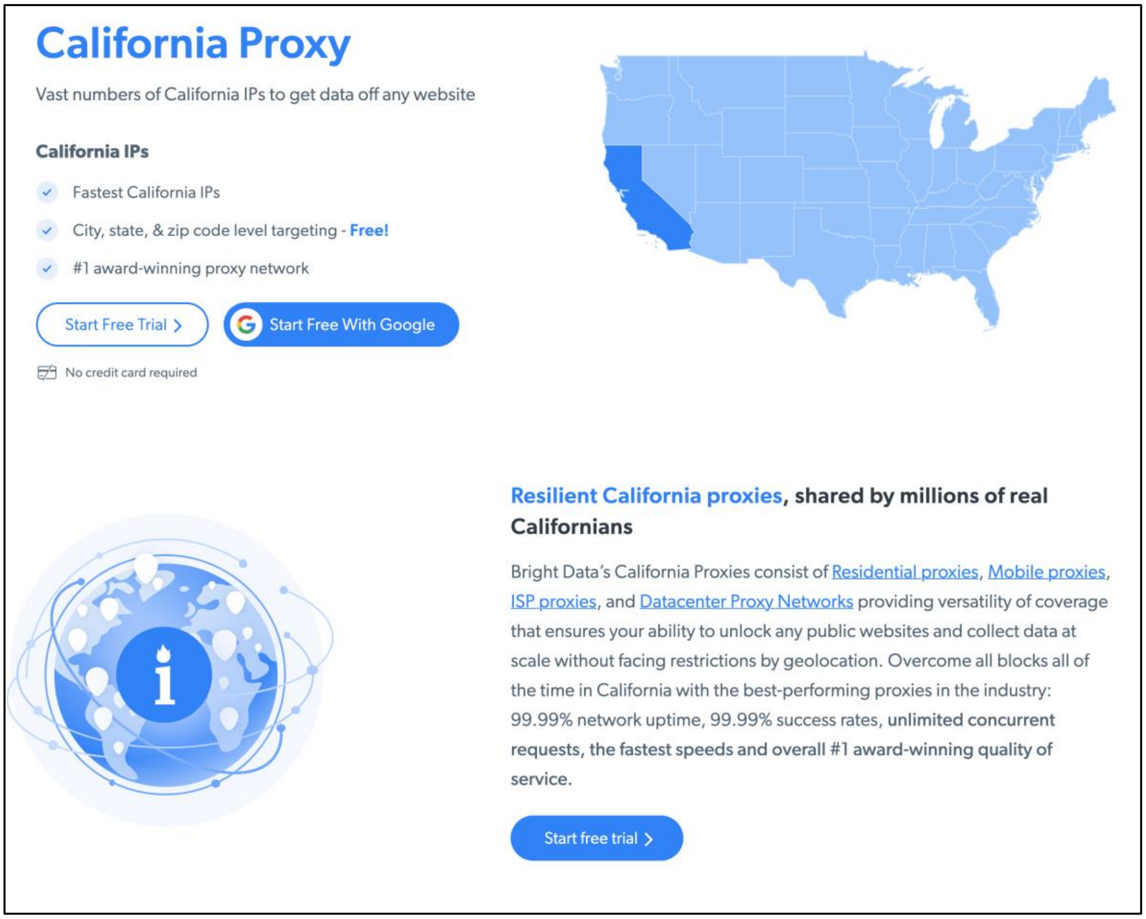
Our mission is to innovate and leverage AI as the ethical, responsible and transparent leader of the public web data industry, while shaping the future of public data accessibility.

- ✓ Most trusted Web Data Platform
- ✓ Founded 2014
- ✓ 450+ team members
- ✓ 20,000+ customers
- ✓ 5,500+ patents
- ✓ HQ in Netanya, IL
- ✓ Offices in NYC & San Francisco
- ✓ Privately held by EMK Capital

Bright Data, *Media Center*, https://perma.cc/Q9YW-V3C4.

~~13.~~17. Defendant ~~has~~ also ~~specifically targeted~~targets its products at the California market. For ~~instance~~example, Defendant's interactive website, through which California residents can purchase Defendant's scraping tools and scraped data sets, offers a "~~Superior~~ California Proxy" product that promises "[v]ast numbers of California IPs to get data off any website."

Figure ~~3~~4: Screenshot from Bright Data's website on ~~November 14, 2023~~August 11, 2024



California Proxy

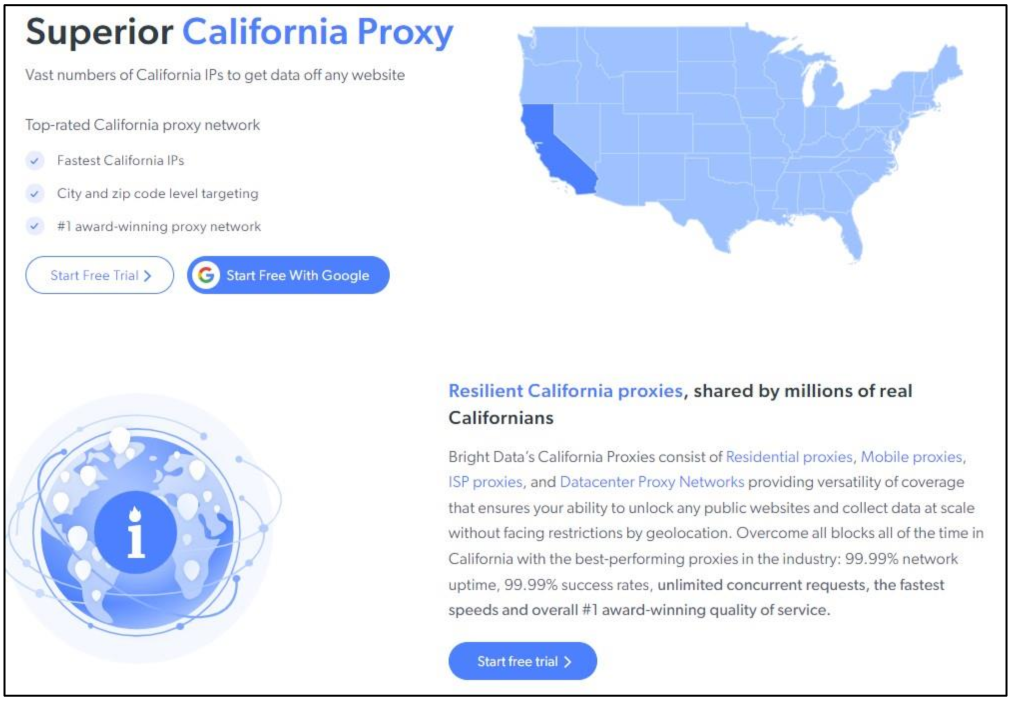
Vast numbers of California IPs to get data off any website

California IPs

- ✓ Fastest California IPs
- ✓ City, state, & zip code level targeting - Free!
- ✓ #1 award-winning proxy network

Resilient California proxies, shared by millions of real Californians

Bright Data's California Proxies consist of Residential proxies, Mobile proxies, ISP proxies, and Datacenter Proxy Networks providing versatility of coverage that ensures your ability to unlock any public websites and collect data at scale without facing restrictions by geolocation. Overcome all blocks all of the time in California with the best-performing proxies in the industry: 99.99% network uptime, 99.99% success rates, unlimited concurrent requests, the fastest speeds and overall #1 award-winning quality of service.



Superior California Proxy

Vast numbers of California IPs to get data off any website

Top-rated California proxy network

- ✓ Fastest California IPs
- ✓ City and zip code level targeting
- ✓ #1 award-winning proxy network

Resilient California proxies, shared by millions of real Californians

Bright Data's California Proxies consist of Residential proxies, Mobile proxies, ISP proxies, and Datacenter Proxy Networks providing versatility of coverage that ensures your ability to unlock any public websites and collect data at scale without facing restrictions by geolocation. Overcome all blocks all of the time in California with the best-performing proxies in the industry: 99.99% network uptime, 99.99% success rates, unlimited concurrent requests, the fastest speeds and overall #1 award-winning quality of service.

~~*See* Exh. D.~~ These proxy IP addresses are designed to evade usage restrictions and anti-scraping technology, such as those implemented by X. In fact, Defendant ~~specifically~~ advertises that its California proxies allow users to “[o]vercome all blocks all of the time in California.” Bright Data, *California Proxy*, https://perma.cc/L7GE-CXUP.

~~14.~~18. On information and belief, Defendant has sold its scraping tools, scraped data sets, and IP proxies to X users, including X users in California, and has scraped data from X Corp.’s servers in California.

~~15.~~19. Venue is proper in this district under 28 U.S.C. § 1391, because a substantial part of the events or omissions giving rise to the claims occurred in this judicial district. During all relevant times, Defendant repeatedly, knowingly, and intentionally targeted its wrongful acts at X Corp., which has its principal place of business in this district. Defendant also, on information and belief, sold its scraping tools, scraped data sets, and IP ~~Proxies~~proxies to residents of this district, including through Defendant’s sales office located in this district and employees located in this district.

~~16.~~20. Pursuant to Civil L.R. 3-2(d), this case may be assigned to either the San Francisco or Oakland division because X Corp. is located in San Francisco County.

FACTUAL ALLEGATIONS

A. - X Corp.’s Platform and Terms of Service

~~17.~~21. Plaintiff X Corp. owns and operates the social media platform X, accessible through twitter.com, X.com, and various mobile and online applications.

22. X serves multiple audiences, including users and developers. For its users, X contributes to the public conversation. X serves its users content on a variety of interests and topics and allows its users to engage with and post their own content. For its developers, X provides automated and programmatic access to the content on X so that businesses, researchers, and developers have real-time access to the global conversation subject to safeguards that protect the platform’s integrity and the X’s user experience. The following paragraphs address each in turn.

1. X’s Services for Users & Terms of Service

~~18.~~23. The X user platform has hundreds of millions of active users worldwide. More than 23 million X accounts have been registered from California.

~~19.~~24. ~~X Corp.~~To create a forum for a global conversation about “what’s happening,” X allows its registered users to post and share content, including written comments, images and videos, known as “posts~~,~~,” and to share, like, and comment on other users’ posts.

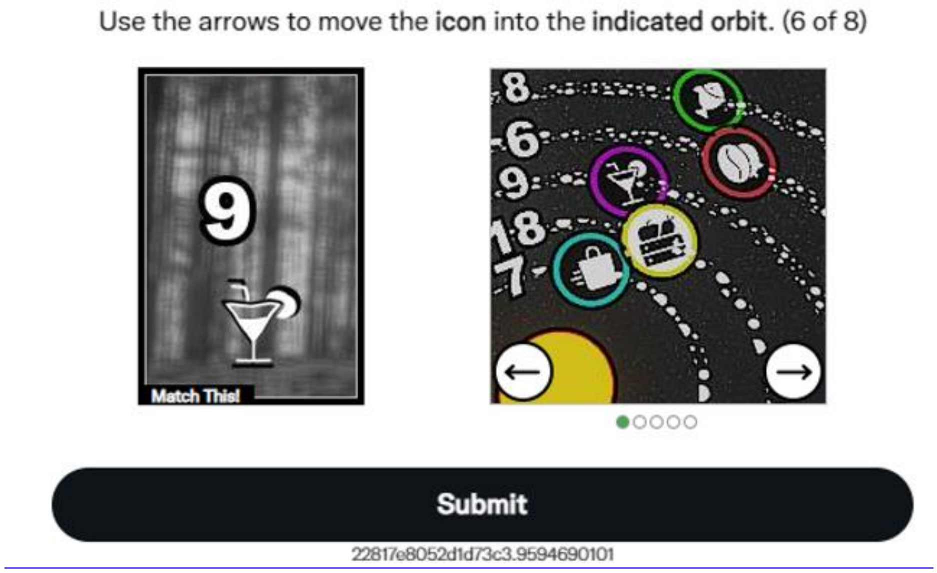
25. To post content on X or to re-post, like, or otherwise interact with posts by others, users must register for an account and log in to that account. As of July 2023, X has also strictly limited the access that individuals (or bots) have when not logged into a registered account. To gain full access to the platform, an individual must be signed into an account.

~~20.~~26. If an unregistered user attempts to access the platform through an app, that user is prompted to create an account (which requires agreement to the Terms) and cannot use the app until they do so. If an unregistered user visits the X homepage at x.com or twitter.com, the user is invited to create an account or sign in.

~~21.~~27. To register for an account, users must provide their name, phone number or email address, and date of birth. To prevent automated services from ~~registering for~~ creating numerous fake accounts, X ~~Corp. requires potential account holders to complete~~imposes a “CAPTCHA” ~~fraud-detection~~ process to ~~determine whether~~ensure that a human (rather than an automated process) is creating the ~~user is human. X Corp.~~account. CAPTCHA stands for “Completely Automated Public Turing test to tell Computers and Humans Apart.” X then verifies registrants through email or phone confirmation.

~~All users who register for a X account, and/or view the X website or application agree to a binding contract with X Corp. as outlined in X Corp.’s User Agreement, which is comprised of the Terms of Service, Privacy Policy, and the Rules and Policies (collectively the “Terms”).~~

28. ~~X Corp.’s~~A picture of X’s CAPTCHA is below. To finish the registration, the prospective user must supply the information requested by the CAPTCHA:



Use the arrows to move the icon into the indicated orbit. (6 of 8)

~~22.~~29. The registration process also requires express agreement to X Corp.'s Terms. The Terms state that a user may not "access, tamper with, or use non-public areas of the Services, our computer systems, or the technical delivery systems of our providers" or "breach or circumvent any security or authorization measures."

~~23.~~30. X Corp.'s Terms also state a user may not "access or search or attempt to access or search the Services by any means (automated or otherwise) other than through our currently available, published interfaces that are provided by us (and only pursuant to the applicable terms and conditions), unless you have been specifically allowed to do so in a separate agreement~~ with us~~."

~~24.~~31. In addition, X Corp.'s Terms specifically state that "crawling or scraping the Services in any form, for any purpose without our prior written consent is expressly prohibited."

~~25.~~32. Under the Terms, users may not "forge any TCP/IP packet header or any part of the header information in any email or posting, or in any way use the Services to send altered, deceptive or false source-identifying information."

~~26.~~33. Users are also prohibited under the Terms from any conduct that would "interfere with, or disrupt, (or attempt to do so), the access of any user, host or network, including … overloading, flooding, spamming … or by scripting the creation of Content in such a manner as to interfere with or create an undue burden on the Services."

~~27.~~34. The Terms also incorporate by reference X Corp.'s Platform Manipulation and Spam Policy (the "Policy"), which ~~specifically~~ prohibits "coordinated harmful activity that encourages or promotes behavior which violates [X Corp.'s] Rules." The Policy also prohibits "leveraging X's open source code to circumvent remediations or platform defenses."

~~28.~~35. The Terms prohibit selling any content collected from the platform. Users may not "reproduce, modify, create derivative works, distribute, sell, transfer, publicly display, publicly perform, transmit, or otherwise use the Services or Content on the Services" unless otherwise authorized by the Terms or a developer agreement.

~~The Terms also expressly state that it is "a violation of these Terms to facilitate or assist others in violating these Terms, including by distributing products or services that enable or encourage violation of these Terms."~~

~~Advertisers on the X platform are also subject to X Corp.'s Ads Policies, which expressly state that advertisers must follow the Terms and all X Corp. Rules.~~

36. ~~For developers who wish to retrieve or analyze X Corp.'s data,~~ The Privacy Policy allows X users to choose their privacy settings from a menu of options giving them the ability to withdraw their consent for data sharing or to choose what content is publicly shared. X users may choose to limit access to their content so that only specified individuals – not all X users – may view their content. They may adjust their privacy settings whenever they wish.

37. X also gives its users the option to delete their posts or their interactions with other users' posts.

38. X also allows users subject to the jurisdiction of, for example, the California Consumer Privacy Act or the European Union's General Data Protection Regulation, to exercise their privacy rights under that act and regulation, including by making certain deletion requests.

X's Services for Developers &~~X Corp. offers specialized access to its Application Programming Interfaces ("APIs") through a tiered subscription service.~~

2. ~~***X Corp.'s***~~ ***Developer Agreement*** ~~***also limits the access of developers to X Corp.'s content.***~~

39. The ~~Agreement instructs~~data available on X – including the way X Corp. has organized it on its platform – has tremendous value to many parties. The data includes both user-

generated content (like user posts and profiles) and non-user-generated content (like follower lists and other information about the relationships between users). The latter type of data typically reflects insufficient originality to warrant copyright protection. But just as important as the data available on the X platform is how the platform organizes it. There are more than 500 million posts on X per day. Determining which posts to show to which users at which times, and on which parts of the website or app, is a significant part of the platform's value. Beyond individual user posts, the *aggregate* data set across X's user base – amalgamating different posts or user interactions to discern macro trends – also delivers unique value to advertisers and developers. X users may have a copyright interest in the individual content they post to X (though they give X a broad license to that content), but they have no copyright interest in much of the most valuable data available on X – including the non-user generated content, the organization of that content, and the aggregate data across X's platform.

40. Businesses, researchers, and others seek the data available on X for a variety of purposes, including measuring user sentiment toward various products or events, gauging market reactions to current events, more effectively tailoring advertisements, and more.

41. The data is primarily valuable when it can be analyzed at scale to discern trends, gauge market reactions, or refine ad campaigns. Data scrapers are not primarily interested in individual posts in which the user may hold a copyright. That is because any individual piece of content in isolation – like a user-generated video – has limited value to businesses looking to discern broader trends in the data (though, of course, users can and do enjoy individual posts). For that reason, scrapers like Bright Data typically target everything, siphoning off data from X at massive scale.

42. X has copyrights to its website and app, which Bright Data accesses when it scrapes data in circumvention of X's technological safeguards. By contrast, this complaint disclaims any exclusive copyright in X's users' posts, and X does not seek to enforce any copyright its users retain in their own creative works. Indeed, X acknowledges that its users retain the right to sell or license their posts to others, including Bright Data, just as they retain the right to exclude others from exploiting those posts. But Bright Data may not scrape those posts from X's platform – without consent from X or its users – and sell them as part of the massive data packages it markets.

43. X has historically made APIs available to developers ~~that they may "not exceed or circumvent rate limits, or any other limitations or restrictions described in this Policy or your agreement~~ to allow them to programmatically interact with ~~Twitter, listed on the~~ X in a controlled, secure environment. This is known as X's Developer ~~Site, or communicated to you~~Platform.

44. X launched the current iteration of its Developer Platform in February 2023. It has four tiers of access: Free, Basic, Pro, and Enterprise.

45. The Free tier does not allow developers to obtain data from the X platform. This tier allows developers only to post to X and test the API.

46. The Basic and Pro tiers are for hobbyists, researchers, and smaller businesses testing out the API. These tiers cost $100 per month (for Basic) or $5,000 per month (for Pro). Both levels impose rate limits that limit the amount of data that can be obtained over a set amount of time (e.g., approximately 900 posts per 15 minutes for Pro), and likewise impose monthly caps on data that can be obtained (e.g., one million posts per month for Pro).

47. The Enterprise tier is for business and scaled commercial projects that need access to data on X's platform at great scale and in real-time. This tier is ideal for large businesses tracking consumer sentiment, financial companies analyzing market trends, or advertisers. The usage and rate limits are far higher. The entry level fees start at [redacted] and rise depending on the usage and rate limits, but more importantly, the use case and industry potentially served.

~~29.~~48. Because of the large amount of available data, X requires that developers seeking access to the Enterprise tier submit their proposed use cases so that X can ensure those uses are consistent with a healthy platform and do not undermine the interests of X's users. X has rejected several requests for access to Enterprise tier – forgoing the associated revenue – where the proposed use cases would have been bad for X users, including by ~~Twitter."~~infringing on X users' privacy.

49. No API tier allows developers unrestricted access to all data on X. For example, X does not allow any API access originating from IPs in certain high-risk jurisdictions, such as those that might promote terrorism. Instead, X's API blacklists IPs from those geographic locations, with the goal of preventing malign actors from gaining unfettered access to X's data.

X also does not allow API access to any governments, to reduce the risk of government surveillance of X's users.

50. Not all X user content, including some X user content which is publicly accessible on the user side of the X platform, is available through the API. This includes, for example, certain information related to user's specific geographical locations and other information detailed below.

51. X Corp. requires each developer using its API – at any tier – to agree to X's Developer Agreement (last updated Nov. 14, 2023), https://perma.cc/2YCC-E6E7. That agreement imposes the following restrictions:

a. X developers must obtain X user consent before sharing an individual user's content to promote a product or service, and before storing or sharing non-public or confidential X user information.

b. X developers may not circumvent user blocking or account protections.

c. X developers must delete from their databases any content that is deleted on X, whether deleted by a user or deleted by X for violations of the Terms or applicable laws – for example, revenge-porn content. As part of this, developers must delete X user content after an X user deactivates or deletes their X account.

d. X developers must modify any data modified on X, including when content is made private or deleted.

e. X developers must not collect X user geodata on a standalone basis, barring them from engaging in user tracking, activity heat maps, or similar activities.

f. X developers must not use the X data to create spam, X bot accounts, or automate processes on the X user side such as bulk X user following.

g. X developers must not use X data to infer certain protected characteristics of X users which X does not share with developers even if an X user publicly posts this content on X's user platform. These protected characteristics include: health (including pregnancy), negative financial status or condition, political affiliations or belief, racial or ethnic origin, religious or philosophical affiliation or beliefs, sex life or sexual orientation, trade union membership, and whether the user has actually or is alleged to have committed a crime.

h. X developers must comply with X user requests which X forwards to them under applicable privacy laws, including the California Consumer Privacy Act and General Data Protection Regulation.

i. X developers must not attempt to match X content, usernames, or accounts with a person, household, device, browser, or other off-X identifier without the user's express consent, unless the information was provided by the user or is otherwise publicly available (i.e., for public figures).

j. X developers must not use acquired data for tracking or targeting sensitive groups, such as political activists or dissidents, performing background checks or personal vetting, credit or insurance risk analysis, individual profiling or psychographic segmentation, or the development of facial recognition software.

3. *Balancing User, Developer, and Advertiser Interests*

52. To balance the complex, interrelated nature of its multi-sided platform business and simultaneously provide a forum for global public conversation on "what's happening," X must be able to credibly enforce its Terms on its users, developers, and advertisers. For example, businesses (including advertisers) may want greater access to data on X for the purpose of targeting specific consumers, but some of X's users may be sensitive to such practice. X Corp.'s Terms restrict such targeting to protect its users' interests. To take another example, some businesses may want unrestricted API access so they can send solicitations to users by direct message or offer to sell increased "follower counts," but X bars such practices because they degrade the quality of X for almost all users. And finally, some governments may seek unrestricted API access to influence or monitor public opinion – again, a practice that disserves X's users.

53. X Corp. ~~Background on~~ can police these problems only if it can credibly enforce its policies, including specifically on programmatic or API access to X. For example, over a recent six-month period, X has enforced its platform manipulation and spam policies by:

a. Suspending 169,396 accounts and removing 15,275 instances of conduct for violations of X's policies against sharing improper impersonation;

b. Suspending 2,563 accounts and removing 62,537 instances of conduct for violations of X's policies against sharing personally identifiable information; and

c. Suspending 119,508 accounts and removing 571,902 instances of conduct for violations of X's policies on illegal and regulated goods.

54. X also maintains an active civic integrity policy.[1] This policy prohibits users from posting false and misleading information about how to participate in a civic process, including elections. For example, users may not post information intended to mislead or intimidate voters to dissuade them from participating in elections.

55. X's civic integrity policy also bars users from posting from accounts with deceptive identities. In February 2021, X disclosed it removed 373 accounts and related instances of content attributed to state-linked information operations originating from Iran, Armenia, and Russia. X again disclosed in December 2021 that it removed 3,465 accounts connected to state-linked information operations from six distinct jurisdictions: Mexico, the People's Republic of China (PRC), Russia, Tanzania, Uganda, and Venezuela. Every account and piece of content associated with these operations was permanently removed from the X platform.

56. X further maintains policies against the use of deceptive marketing or misrepresentative business practices,[2] as well as the advertising of certain high-risk financial products and certain content related to cryptocurrencies.[3]

[1] X Help Center, *Civic integrity policy* (Aug. 2023), https://help.twitter.com/en/rules-and-policies/election-integrity-policy.

[2] X Business, *Deceptive & Fraudulent Content Policy*, https://perma.cc/XAU6-S648.

[3] X Business, *Financial products and services*, https://business.x.com/en/help/ads-policies/ads-content-policies/financial-services.html.

57. The availability of unrestricted third-party programmatic access (including scraping data) weakens X's ability to effectively prevent and deter market manipulation, scams, and fraud, increasing the risk of harm to consumers.[4] This is because the most effective way for X to police such abuses is by having a business relationship with and approving the use cases of those third parties given access to the X API. The goal of data scrapers, by contrast, is to conceal their conduct from X and sell the data they scrape without restriction.

B. X's Battle Against Unapproved Programmatic Access Including Data Scraping

58. Many malign actors neither want to pay for API access to X nor want to comply with X's Developer Agreement. Indeed, many want programmatic access to X for purposes that X's API disallows – such as spamming X users with solicitations or disinformation or scraping vast swaths of data without regard to user privacy. Because these actors cannot acquire X's data through legitimate channels, they turn to services like Bright Data to get the job done.

~~30.~~59. One form of programmatic access – data scraping – is particularly harmful to X. Scraping is the process of using automated means to collect content or data from a website~~.~~ or app. The process involves ~~making~~ using an automated process to make a request to a website's server or app, downloading the results, and parsing them to extract the desired data. Data scrapers typically send large volumes – in the millions or even billions – of these requests, taxing the capacity of servers and diminishing the experience for legitimate users. Data scraping is also an end-run around X Corp.'s restrictions on API access that protect the user experience on X.

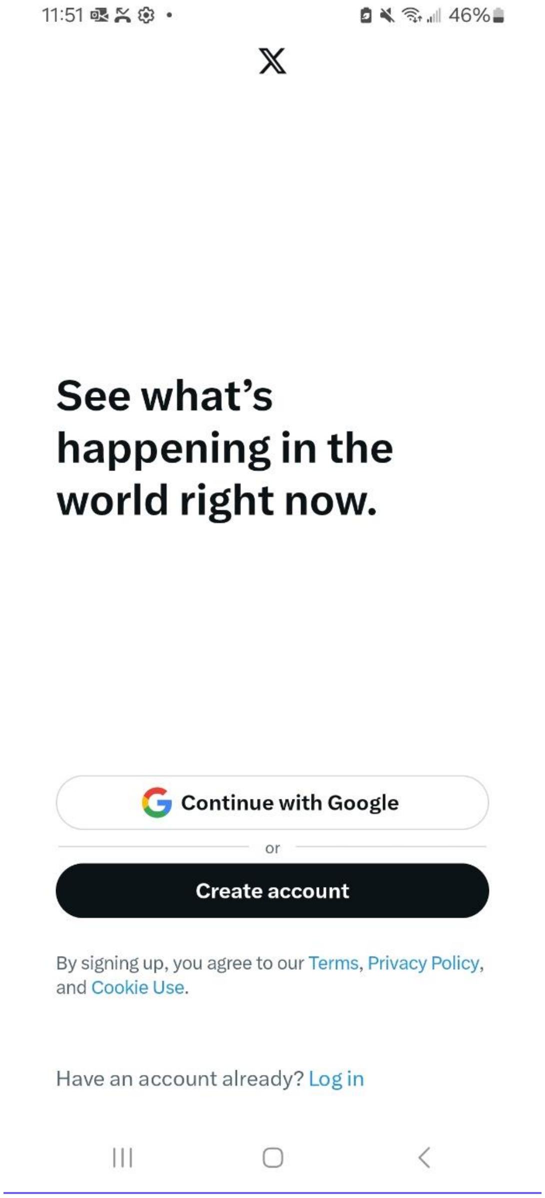
60. X Corp. utilizes a variety of technological measures to detect and prevent automated systems – colloquially known as "bots" – from scraping data from its platform~~, including industry standard automation prevention techniques~~:

1. User Login Requirements

[4] *See, e.g.*, Press Release, SEC, *SEC Charges Eight Social Media Influencers in $100 Million Stock Manipulation Scheme Promoted on Discord and Twitter* (Dec. 14, 2022) https://www.sec.gov/news/press-release/2022-221; Press Release, U.S. Att'y's Off., N.D. Cal., *Scottish Citizen Indicted For Twitter-Based Stock Manipulation Scheme* (Nov. 5, 2015), https://perma.cc/8NPH-GN3W.

61. Since July 2023, X does not make most of its content available to the public without logging in through a registered account and agreeing to the Terms.

62. X's apps cannot be used without logging in. Whenever the app is opened without being logged in, the app prompts the user to login or create an account and cannot be used further until the user does so.

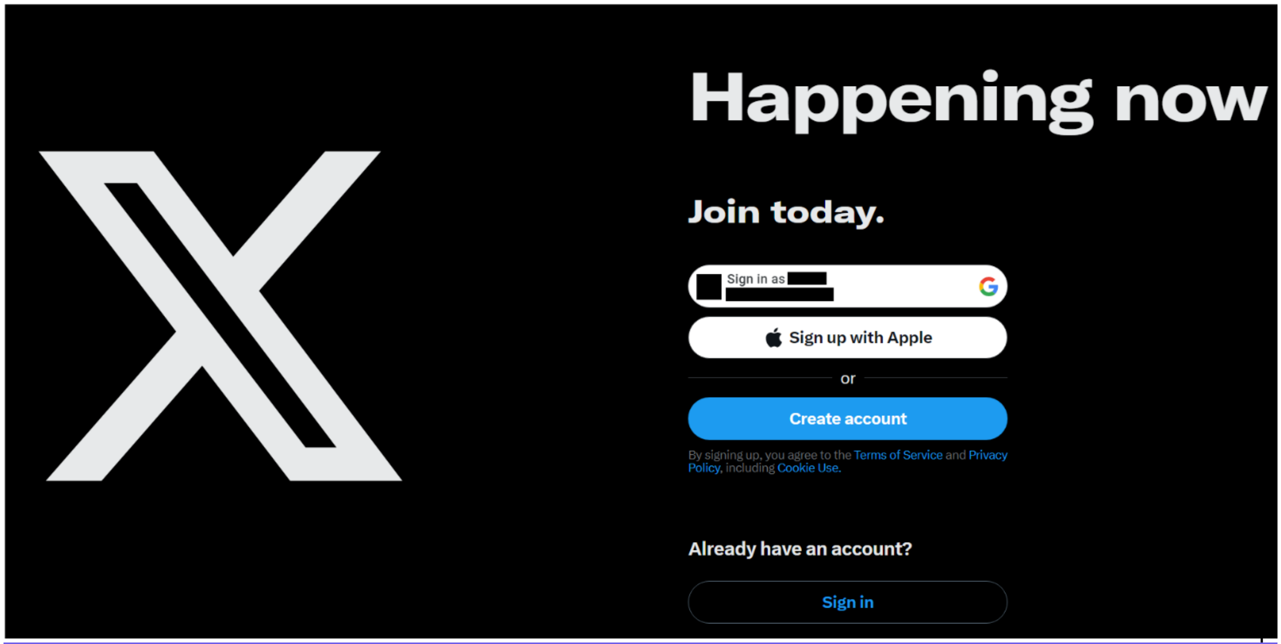


See what's happening in the world right now.

By signing up, you agree to our Terms, Privacy Policy, and Cookie Use.

Have an account already? Log in

63. The homepages of X's website, x.com and twitter.com, also cannot be used without logging in. If a user is not logged in, the homepage prompts the user either to login or create an account and displays no X content until the user does so.



Happening now

Join today.

[redacted] Sign in as [redacted]

Sign up with Apple

or

Create account

By signing up, you agree to the Terms of Service and Privacy Policy, including Cookie Use.

Already have an account?

Sign in

64. X makes a limited amount of content available to individuals who are not logged in. Examples include users who navigate to X's website from another source, such as ~~CAPTCHAs, user identification and IP~~ a search engine or an X post that is embedded within another website. X also allows Google to "crawl" certain posts on its website to index certain posts so that they are shown in Google's search results. Not all posts on X are accessible from another source.

65. If non-logged-in individuals access a post on X in this manner, they can view that specific post and certain limited information about the post (e.g., number of likes, replies, and reposts). But their visibility is limited. The individual cannot see any replies to the post or the identities of who liked or reposted that content – thus disabling a key platform feature.

66. A non-logged-in individual accessing X in this manner also has a limited ability to access other content on X. The individual can access the profile of the person who posted the content but can see only a curated list of other posts from that individual (not all posts) and cannot see that user's list of followed accounts or followers. The non-logged-in individual can also click on other linked posts, which show those posts subject to the same restrictions above.

67. When users are logged out, X circumscribes their ability to continue accessing different profiles or linked posts. After several clicks on different profiles or posts, X prompts the user who is not logged in to either login or create an account. At that point, the user may no longer

click on additional profiles or posts until a period of time has passed. Thus, the ability to access content when not logged in is both rate limited and qualitatively restricted.

~~31.~~68. As set forth above, X enforces these rate limits by IP address and through other anomaly~~-~~-detection tools. To access X's website, a user must provide their IP address, supply identifying markers associated with their web browser, and also make requests about how they would like to interact with the X platform. X's technological measures analyze this information to determine whether the request comes from a legitimate human user or an unauthorized automated process.

69. ~~X Corp.'s registration process~~The only feasible means of data scraping the limited content that is accessible to non-logged-in users is by circumventing X's IP rate limits and anomaly-detection tools.

2. Account Creation Restrictions

~~32.~~70. To ensure that malicious actors cannot automatically create numerous accounts for impermissible purposes, including data scraping, X Corp. requires potential registrants to pass a CAPTCHA ~~and provide~~ by inputting the required information; to enter a valid phone number or email address~~. Prior to creating the user's account,~~ ; and to supply ~~X Corp. sends~~ a verification code X sends to ~~the~~that email or phone number. ~~Potential users must enter the verification code to create an account.~~

~~X Corp. also employs rate limits that cap the number of posts that may be viewed by registered users and those who access the platform without an account. Developers who use the X API are also capped in the number of posts they may post to, or pull from, the platform based on their subscription level~~**3. Rate Limits**

71. X further limits user access to X content even once logged in, in an effort to fight data scraping through logged-in accounts. In July 2023, X imposed limits for the number of posts that a single account can view in a day: 500 for new unverified accounts, 1,000 for unverified accounts in regular use, and 10,000 for X premium subscribers.

72. These reasonable limits for human use make mass data scraping impossible. Accordingly, the only way to data scrape X is to circumvent these rate limits by creating large

numbers of bot accounts. And to create large amounts of bot accounts, the data scraper must circumvent X's account-creation restrictions.

4. Anomaly Detection Tools

~~33.1.~~ X also employs ~~.~~

73. ~~X Corp.~~ ~~also utilizes~~ anomaly detection tools to detect attempted use of many accounts for data scraping or other impermissible purposes, including at the account-creation stage. These tools analyze the IP address being used, certain identifying information about the web browser being used, and ~~block~~ characteristics of the type of use (such as patterns that indicate automated ~~software that is attempting to~~ access).

74. To access X's service, users must provide the information relied upon by these anomaly detection tools, including the user's IP address, identifying information for the web browser, and the actions the user wishes to perform on X.

5. Robots.txt

75. In July 2023, X Corp. modified its robots.txt instructions to prohibit all forms of automated access to X's website except for Google's web crawler, which is used to index websites for placement on Google's search engine. X Corp.'s robots.txt instructions explicitly forbid any scraping of its website by others, like Bright Data or Bright Data's customers.

76. Robots.txt is a file placed on a website that instructs automated bots which content (if any) they are allowed to access. Robots.txt is an industry-standard tool used by website operators since 1999.

77. Bright Data, as one of the largest (if not the largest) providers of data-scraping tools, is familiar with the existence and meaning of instructions in a website's robot.txt file.

78. All legitimate operators of automated bots (such as search engines like Google) comply with the instructions in a robots.txt file. Bright Data, by contrast, willfully ignores them. Bright Data claims it will force compliance with robots.txt directives for its "immediate access" mode in its "Residential Proxy Network" because that is "respectful and compliant." But it also makes clear that it offers its customers "full . . . access" – that is, not constrained by robots.txt directives – so long as the customer provides Bright Data with basic contact information. *See* Bright Data, *Residential Proxy Network Policy*, https://perma.cc/B945-FX6U. This process is a

wink-and-a-nod that Bright Data will ignore robots.txt directives in facilitating full-scale access to data on the platforms it targets.

* * *

79. Because of the combined technological measures X erects to protect its platform from automated access, the only way to data scrape X at the scale Bright Data promises is by: (1) circumventing X's account-creation restrictions like CAPTCHA to automatically open legions of fake accounts; (2) circumventing X's use of IP address monitoring to permit these fake accounts to bombard X's servers at rates otherwise barred by X's limits; and (3) circumventing X's other anomaly-detection tools by having an automated process submit requests to appear as if it were a regular human user.

C. The Harm X Suffers from Bright Data-Enabled Automated Access

80. X has expended vast sums of money developing a service that Bright Data appropriates for illicit ends. X's service is far more sophisticated than one that simply displays user posts. On average, 250 million registered users login to X daily, and 550 million unregistered users visit monthly. On average, there are about 500 million posts per day. Displaying, organizing, and managing all this content requires both money and expertise. For example, X's Recommendation Algorithm must decide hundreds of millions of times per day which of billions of posts to display to its users, and that must happen all in a matter of milliseconds every time. That process involves multiple complex steps to filter posts down to a smaller set of candidate posts, which are then ranked by a continuously trained neural network.[5]

81. The infrastructure needed to power that Recommendation Algorithm (and everything else on X) is massive, spanning two large datacenters owned and operated by X and supported by other cloud providers. That infrastructure processed approximately 400 billion events generating petabyte-scale data every day (as of October 22, 2021) and has cost billions of dollars.[6]

[5] *See* Twitter, *Twitter's Recommendation Algorithm* (Mar. 31, 2023), https://blog.x.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm.

[6] *See* Lu Zhang & Chukwudiuto Malife, *Processing Billions of Events in Real Time at*

82. Data scraping massively taxes X's infrastructure and requires X to spend [redacted] per year to ensure that its service remains responsive to its actual users.

83. [redacted]. And each individual request from a scraper is at least as taxing as an ordinary user request. Most ordinary users are served with a relatively small number of popular posts that X has cached, and the caching makes them relatively efficient for X to serve. When scrapers seek to request *all* posts, by contrast, it forces X to bypass its cache and serve posts from more-expensive storage. This process is far more burdensome for X than caching.

84. Data scrapers compound the harm because they do not use X's APIs, which are optimized for high-volume automated requests targeting the precise information a developer seeks (such as the content of relevant posts). Instead, data scrapers access the X platform as a normal user would, which involves X serving *more* information than the data scraper is actually seeking (such as similar recommended posts or additional users to follow), thus imposing more burden on X's servers than would occur if those data scrapers used the API tailored for the information they are seeking.

85. Further, since the filing of this suit in July 2023, the volume of data scraping on X has only *increased* despite further attempts by X to stop it. As Bright Data touts on its website, data scraping (euphemistically called the "alternative data market") is growing at a 46.5% annual rate, due in no small part to Bright Data.

86. X monitors platform usage to identify uses that are likely inauthentic and not attributable to human users. The statistics identified below reflect data gleaned from that monitoring and represent actual burdens on X's server capacity, at a minimum.

~~34.~~87. On average, about [redacted] is inauthentic and not attributable to human users. Data scrapers prefer to access X through the web, while most of X's traffic from authentic human users comes via iOS or Android. So the numbers below refer to web traffic. On a normal day, X generally [redacted]

Twitter, X Engineering (Oct. 22, 2021), https://blog.x.com/engineering/en_us/topics/infrastructure/2021/processing-billions-of-events-in-real-time-at-twitter-.

[redacted]. X's monitoring shows that specific types of end points are more likely to be targeted by data scrapers, including up to [redacted] of certain types of access being by data scrapers. Several are identified below. Notably, each of the types of access listed below are available *only* to logged in users. This demonstrates that most data scraping occurs through logged-in accounts that have agreed to X Corp.'s ~~platform.~~Terms (only to blatantly violate them), *not* through logged-out scraping of public data:

a. [redacted] of web requests to view [redacted] are anomalous or inauthentic.

b. [redacted] of the web requests to look up [redacted] are anomalous or inauthentic.

c. [redacted] of the web requests to look up [redacted] are anomalous or inauthentic.

d. [redacted] of web requests to look up [redacted] are anomalous or inauthentic.

e. [redacted] of web [redacted] are anomalous or inauthentic.

88. X has also encountered and investigated specific instances of massive data scraping, which include: In September 2023, [redacted] of all requests for [redacted] were coming from data scrapers. In October 2023, data scrapers were submitting 800,000 requests *per second*. In December 2023, [redacted] of all [redacted] traffic originated from data scrapers.

89. Because X allocates specific server capacity to [redacted], these requests lead to [redacted] for the [redacted] scraping targets. Even when specific servers are overloaded, X's [redacted] attempts to ensure that the system overall is less likely to fail entirely, but even isolated failures lead to [redacted]

90. To avoid the catastrophic failure of X's platform servers and systems due to impermissible data scraping, X obtains an average projected additional [redacted] of over-provisioned headroom for its server load capacity that it needs to accommodate the anomalous bot activity facilitated by data scrapers like Bright Data. This additional server load capacity imposes additional costs ranging from [redacted] each month.

91. If X Corp. did not purchase this additional server load capacity, the X user platform would risk failure and significant degradation of the X user experience. Because X has a reputational and business imperative to be available 24/7 as the platform known for "what's happening" moment to moment, it cannot afford those risks and must pay for additional capacity.

92. X also employs a dedicated team of operational engineers to remedy anomalous access on X, which is constantly evolving as X must understand, assess, and stymie new methods of automated, anomalous, and inauthentic access of its platform. Over the last year, this team responded to and remedied at least [redacted] of known data scraping where the demands on X's server capacity jumped extreme amounts. Of course, this team cannot respond to all scraping, much of which goes unremediated – not because it does not tax X's server capacity but because it is so prevalent.

93. Each of these scraping incidents could occur only through the coordinated use of proxy networks with massive amounts of rotating IP addresses (which Bright Data provides) and the creation of numerous fake accounts automatically through CAPTCHA circumvention (which Bright Data likewise enables).

94. In attempting to stop this activity, X must balance the harms of data scraping with the risk of false positives blocking legitimate access. Data scrapers (like Bright Data or the companies that it enables) succeed by making their activity as indistinguishable from normal user activity as possible – including by use of IP proxies that route traffic through ordinary consumers' devices. Thus, X suffers harm even in attempting to impose technological measures to stop data scrapers. [redacted].

95. Beyond the direct costs posed by the data scraping activity, the continued presence of bots, fake profiles, and data scrapers impacts X's relationship with its users, who participate in X's services on a presumption of trust that X can enforce its policies relating to privacy and user content authenticity, among others. The presence of bots creates risks to privacy and security, subjects users to spam and other unwanted content, and degrades the overall user experience in a way that risks users leaving X. That risks costing X content, data, and other benefits of its large and robust platform.

D. Defendant ~~Has~~and Its Customers Have Agreed to X Corp.'s Terms of Service

~~35.~~96. Defendant has expressly agreed to X Corp.'s Terms and is therefore bound by those Terms.

~~36.~~97. Initially, by using the X platform, Defendant, which is well aware of the Terms, agrees to be bound by them. The Terms specifically state:

> These Terms of Service ("Terms") govern your access to and use of our services, including our various websites, SMS, APIs, email notifications, applications, buttons, widgets, ads, commerce services, and our other covered services (https://help.x.com/rules-and-policies/x-services-and-corporate-affiliates) that link to these Terms (collectively, the "Services"), and any information, text, links, graphics, photos, audio, videos, or other materials or arrangements of materials uploaded, downloaded or appearing on the Services (collectively referred to as "Content"). By using the Services you agree to be bound by these Terms.

~~37.~~98. In addition to agreeing to the Terms by using X services, Defendant, which has maintained a registered account on X (@bright_data) since at least February 2016, expressly accepted and agreed to the Terms when registering its account. Bright Data's X account frequently posts content promoting the company's products and services.

~~38.~~99. Defendant's top executives are also registered X users and expressly agreed to X Corp.'s Terms when registering their accounts, further demonstrating that Bright Data had knowledge of the Terms:

a. Bright Data's CEO, Or Lenchner, has maintained a registered account on X (@orlench) since at least December 2012 and regularly posts from that account.

b. Bright Data's CMO, Yanay Sela, has maintained a registered X account (@yanay_sela) since at least December 2014.

c. Bright Data's Managing Director for North America, Omri Orgad, has maintained a registered X account (@omri_orgad) since at least November 2011.

d. Bright Data's Vice President of Product, Erez Naveh, has maintained a registered X account (@nerez) since at least August 2009.

e. Bright Data's Global Communications Manager, Zachary Keyser, has maintained a registered X account (@KeyserZachary) since at least December 2019.

f. Bright Data's Founder, Ofer Vilenski, has maintained a registered X account (@vilenski) since at least November 2008.

~~39.~~100. On information and belief, several other employees and agents of Defendant involved in Defendant's data-scraping activities are also X account holders, including, by way of example, Artem Shibakov, a Bright Data software engineer who has maintained a registered X account (@ashibakow) since at least February 2013. These account holders have also expressly accepted and agreed to X Corp.'s Terms.

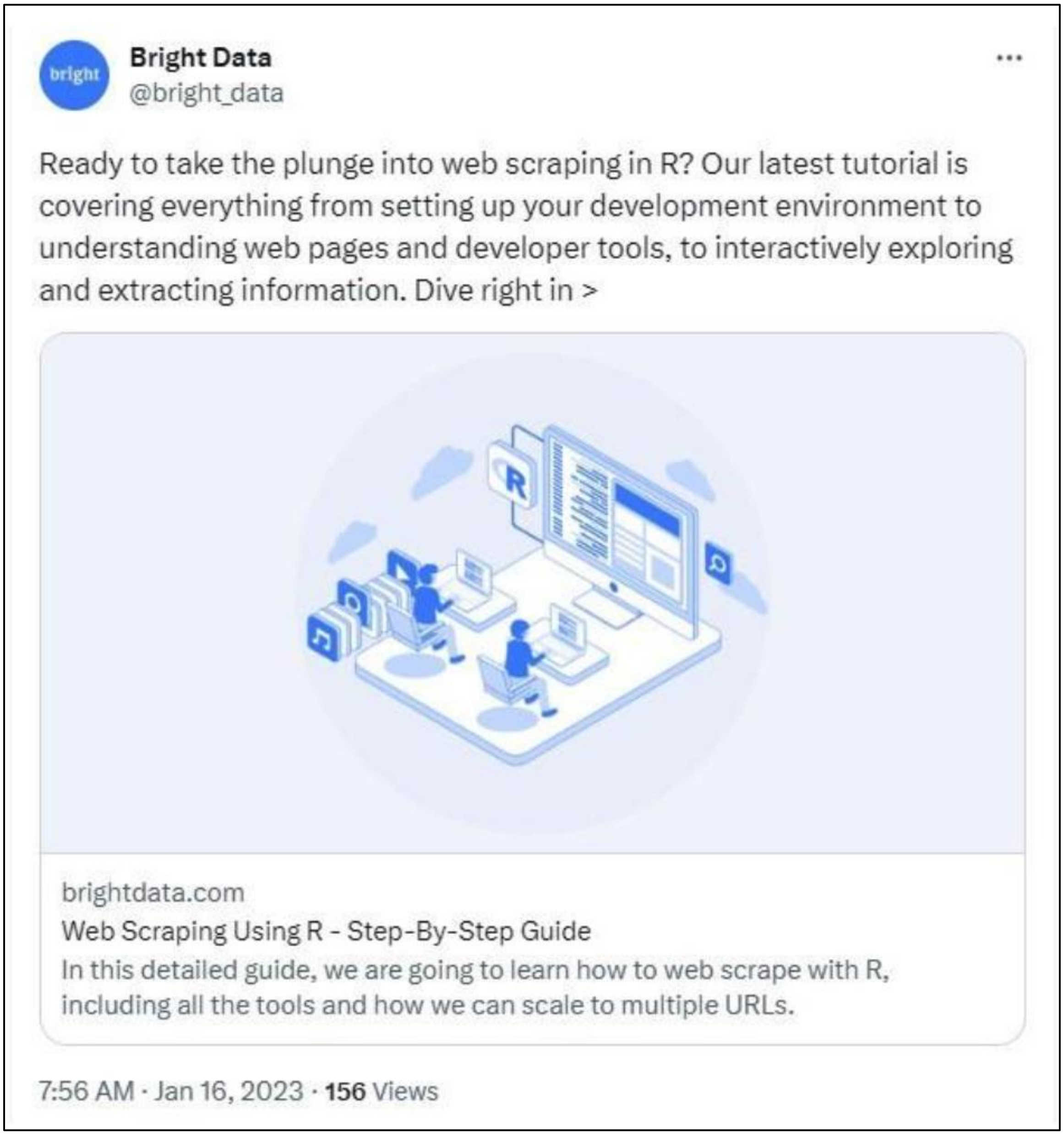
~~40.~~101. Defendant is additionally subject to the Terms as an advertiser on X. Beginning on March 7, 2016, Defendant (then known as Luminati Networks) purchased advertising on the X platform. Defendant purchased additional advertising on X from 2019 to 2021. As stated in X Corp.'s Ad Policies, to which Defendant expressly agreed, all advertisers are bound by the platform's Terms and Rules.

~~41.~~102. Defendant and its executives have repeatedly used these X accounts to discuss and promote their data-scraping products and services, including but not limited to the following posts:

a. On January 1, 2023, Defendant posted a video on X entitled "How to Scrape UNSCRAPABLE data!" which demonstrated how to use Defendant's tools and services for unauthorized data scraping.

b. On January 16, 2023, Defendant encouraged users in a post on X to "take the plunge into web scraping" using a "step-by-step guide" to Defendant's tools and services.

Figure ~~4~~5: Screenshot of Bright Data's X post on July 11, 2023

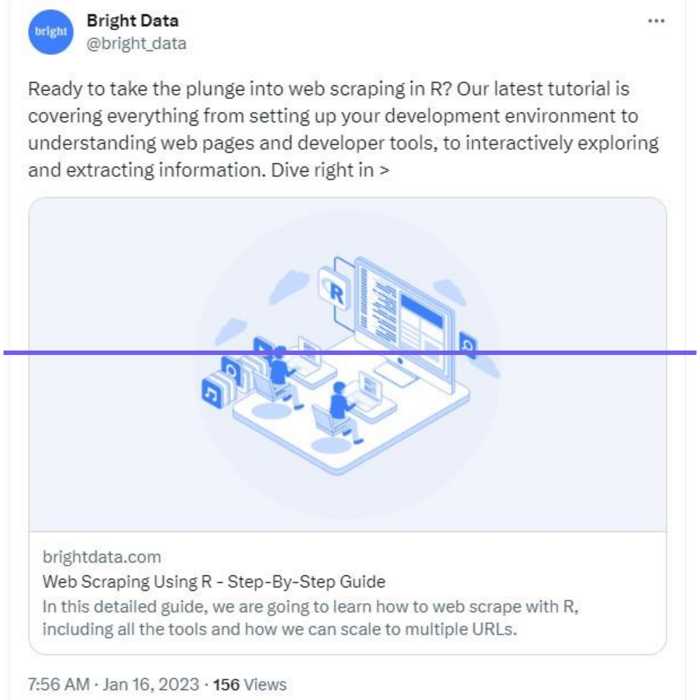


Bright Data
@bright_data

Ready to take the plunge into web scraping in R? Our latest tutorial is covering everything from setting up your development environment to understanding web pages and developer tools, to interactively exploring and extracting information. Dive right in >

brightdata.com
Web Scraping Using R - Step-By-Step Guide
In this detailed guide, we are going to learn how to web scrape with R, including all the tools and how we can scale to multiple URLs.

7:56 AM · Jan 16, 2023 · **156** Views



Bright Data
@bright_data

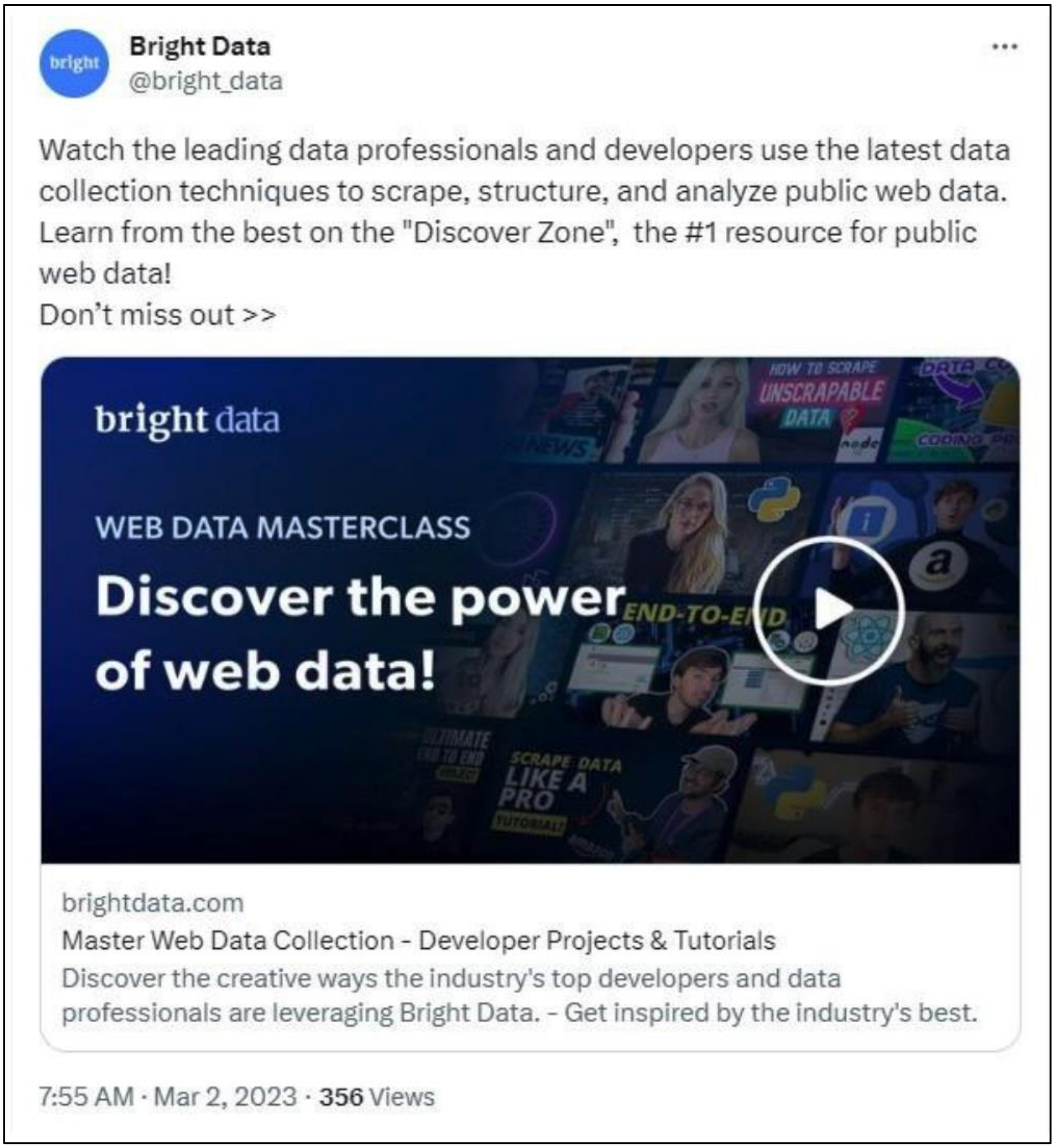
Ready to take the plunge into web scraping in R? Our latest tutorial is covering everything from setting up your development environment to understanding web pages and developer tools, to interactively exploring and extracting information. Dive right in >

brightdata.com
Web Scraping Using R - Step-By-Step Guide
In this detailed guide, we are going to learn how to web scrape with R, including all the tools and how we can scale to multiple URLs.

7:56 AM · Jan 16, 2023 · 156 Views

c. On March 2, 2023, Defendant posted a video on X to a "masterclass" that showed "the latest data collection techniques to scrape, structure, and analyze public web data" using Defendant's tools and services.

Figure ~~5~~6: Screenshot of Bright Data's X post on July 11, 2023



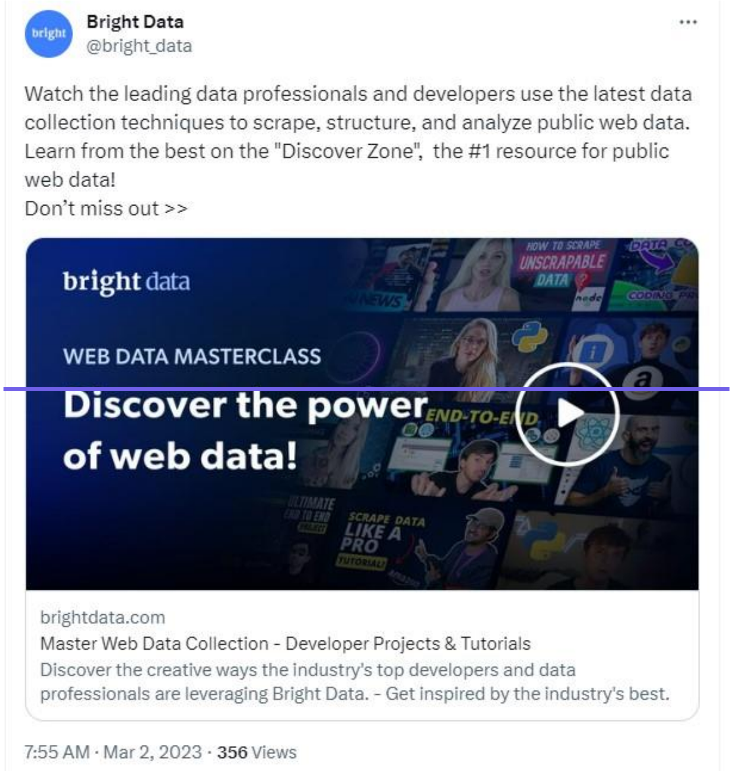
Bright Data
@bright_data

Watch the leading data professionals and developers use the latest data collection techniques to scrape, structure, and analyze public web data. Learn from the best on the "Discover Zone", the #1 resource for public web data!
Don't miss out >>

bright data
WEB DATA MASTERCLASS
Discover the power of web data!
HOW TO SCRAPE UNSCRAPABLE DATA
END-TO-END
SCRAPE DATA LIKE A PRO TUTORIAL

brightdata.com
Master Web Data Collection - Developer Projects & Tutorials
Discover the creative ways the industry's top developers and data professionals are leveraging Bright Data. - Get inspired by the industry's best.

7:55 AM · Mar 2, 2023 · **356** Views



Bright Data
@bright_data

Watch the leading data professionals and developers use the latest data collection techniques to scrape, structure, and analyze public web data. Learn from the best on the "Discover Zone", the #1 resource for public web data!
Don't miss out >>

bright data
WEB DATA MASTERCLASS
Discover the power of web data!
HOW TO SCRAPE UNSCRAPABLE DATA
END-TO-END
SCRAPE DATA LIKE A PRO TUTORIAL

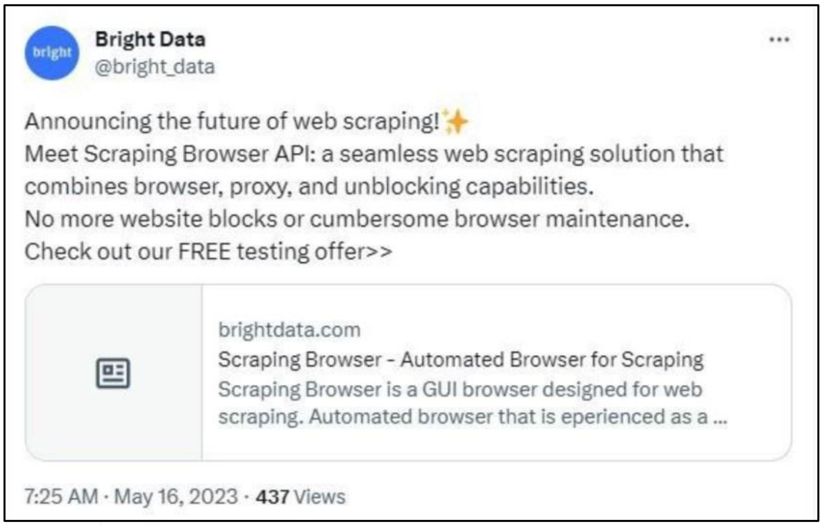
brightdata.com
Master Web Data Collection - Developer Projects & Tutorials
Discover the creative ways the industry's top developers and data professionals are leveraging Bright Data. - Get inspired by the industry's best.

7:55 AM · Mar 2, 2023 · 356 Views

d. On March 23, 2023, Defendant posted ~~a~~and promoted its "Web Unblocker" and its ability to "bypass[] multiple anti-bot solutions" in a post on X.

e. On May 16, 2023, Defendant promoted its "Scraping Browser API: a seamless web scraping solution that combines browser, proxy, and unblocking capabilities" with a link to a "FREE testing offer" in a post on X.

Figure ~~6~~7: Screenshot of Bright Data's X post on July 11, 2023

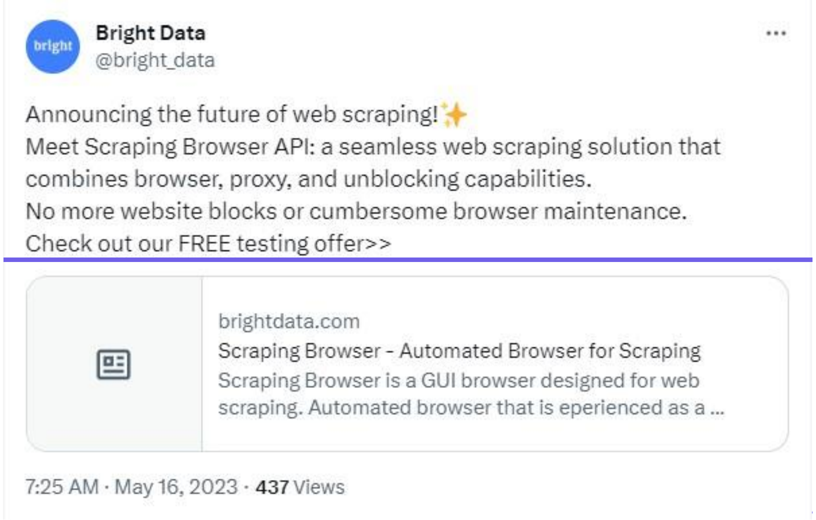


Bright Data
@bright_data

Announcing the future of web scraping!✨
Meet Scraping Browser API: a seamless web scraping solution that combines browser, proxy, and unblocking capabilities.
No more website blocks or cumbersome browser maintenance.
Check out our FREE testing offer>>

brightdata.com
Scraping Browser - Automated Browser for Scraping
Scraping Browser is a GUI browser designed for web scraping. Automated browser that is eperienced as a ...

7:25 AM · May 16, 2023 · 437 Views



Bright Data
@bright_data

Announcing the future of web scraping!✨
Meet Scraping Browser API: a seamless web scraping solution that combines browser, proxy, and unblocking capabilities.
No more website blocks or cumbersome browser maintenance.
Check out our FREE testing offer>>

brightdata.com
Scraping Browser - Automated Browser for Scraping
Scraping Browser is a GUI browser designed for web scraping. Automated browser that is eperienced as a ...

7:25 AM · May 16, 2023 · 437 Views

103. Bright Data's customers are similarly bound by X Corp.'s Terms by use of the X platform, and whenever they sign up for accounts that explicitly require agreement to the Terms.

104. Bright Data is also aware that its customers are bound by X Corp.'s Terms. It is a sophisticated entity that knows those customers are bound by the same Terms to which Bright Data agreed. Further, Bright Data is aware of and complicit in customers' use of Bright Data's tools for the purpose of willfully violating X Corp.'s Terms that prevent data scraping. Bright Data knows that the most commercially valuable uses of data on the X platform require scraping large amounts of data that can only be done through logged-in accounts. Bright Data makes its tools available for the purpose of circumventing X's technological measures designed to prevent that activity.

E. Defendant's Unauthorized Scraping and Sale of Scraping Tools

~~42.~~105. Defendant, per its own admissions, has engaged in widespread scraping of X Corp.'s data, circumventing X Corp.'s technical barriers and violating the Terms to which it agreed. Defendant has also facilitated the scraping of data from X and induced X users to violate ~~X Corp.'s~~the Terms.

~~43.~~106. X Corp. has not granted Defendant permission to scrape data from the X platform.

~~44.~~107. X Corp permits paying ~~customers~~developers to lawfully access certain categories of X data, subject to contractual usage limits and other restrictions designed to protect the X platform and user experience~~.~~, as detailed above. Rather than ~~attempt to~~ lawfully acquire X data through authorized means, Bright Data elected to scrape the data (and enable others to do so), using proxies and other illicit methods to shield its identity and scraping activities.

108. Bright Data markets itself as the leading provider of tools that enable unlawful data scraping of X's platform. Bright Data states that it is the "leading residential proxy provider" in the market, controlling "the best proxy network in the world" and "offer[ing] the best web scraping proxies." *See* Bright Data, *The Top 10 Residential Proxies of 2024*, https://perma.cc/T4MU-MMGQ.

109. Investors have described Bright Data (which used to call itself "Luminati") as "the world's leading enterprise IP proxy network" and "the only mass-scale residential IP proxy network in the world." *See* EMK Capital, *EMK Acquires Luminati – The World's Largest IP Proxy Network Which Brings Transparency to the Internet* (Aug. 10, 2017), https://perma.cc/X4TZ-T9FB.

110. Investors also tout Bright Data as the only product on the market that can evade the technical safeguards of the websites targeted for scraping: "Unlike Bright Data, other IP proxy networks and their customers are prone to being blocked, slowed or spoofed by websites they visit." *See* EMK Capital, *Bright Data*, https://perma.cc/9JM8-MVUS.

111. A 2020 research report described Bright Data as the "[m]arket leader" in proxy network services; another report issued the prior year stated that Bright Data "is the world's largest proxy service." *See* Robert Cavin, *Description – Global Internet Protocol Proxy Network*

Market, Forecast to 2025, Frost & Sullivan (2020), https://perma.cc/CDY6-YJ2K; Business Wire, *Frost & Sullivan Names Luminati the 2019 Global Market Leader in the Enterprise IP Proxy Networks Market* (July 23, 2019), https://www.businesswire.com/news/home/20190723005394/en/Frost-Sullivan-Names-Luminati-the-2019-Global-Market-Leader-in-the-Enterprise-IP-Proxy-Networks-Market.

112. Defendant has not publicly disclosed how it evades X Corp.'s technical safeguards against scraping. Rather, Defendant takes extraordinary steps to conceal its activity from X. Defendant's entire business is built on bypassing the technological measures that X puts in place to stop data scraping – a business that only works if it can evade detection by X.

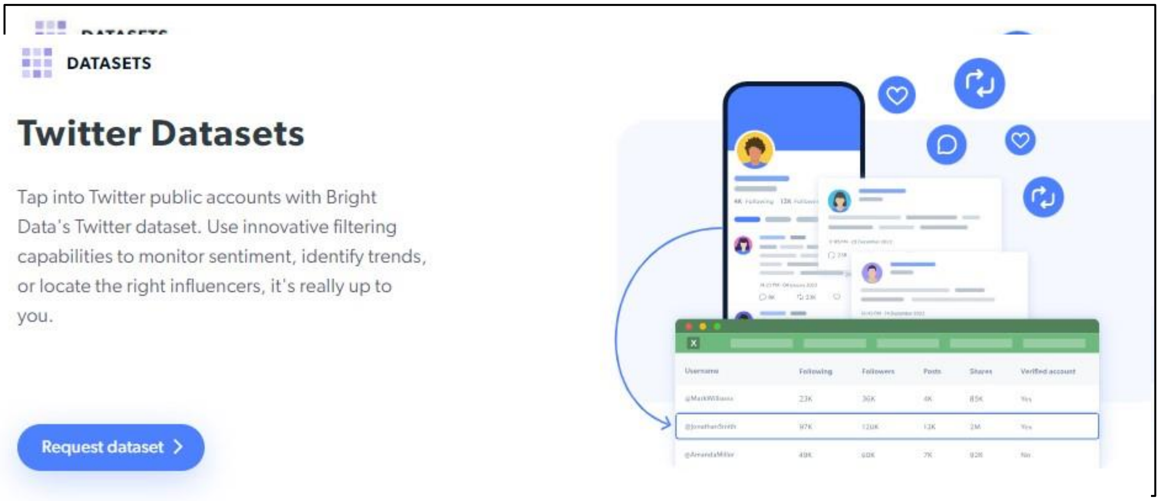
113. Because Bright Data conceals its misconduct, X cannot pinpoint all the instances where Bright Data's customers have used Bright Data's scraping tools. However, because Bright Data is the market leader in providing such tools, because of the increasing level of data scraping that occurs on X, and because of Bright Data's own marketing statements targeting X, it is virtually certain that many major data scrapers of the X platform use Bright Data's tools.

~~45.~~114. Indeed, Defendant's website makes clear that the company itself engages in prohibited scraping of X on an industrial scale and brazenly advertises that Defendant sells tools and services that encourage and enable others to engage in prohibited scraping.

1. Defendant Scrapes and Sells X Corp. Data

~~46.~~115. ~~As seen in Figure 7 below,~~ Defendant offers X Corp.'s data for sale on its website.

Figure ~~7~~8: Screenshot of Bright Data's website on July 10, 2023



DATASETS

Twitter Datasets

Tap into Twitter public accounts with Bright Data's Twitter dataset. Use innovative filtering capabilities to monitor sentiment, identify trends, or locate the right influencers, it's really up to you.

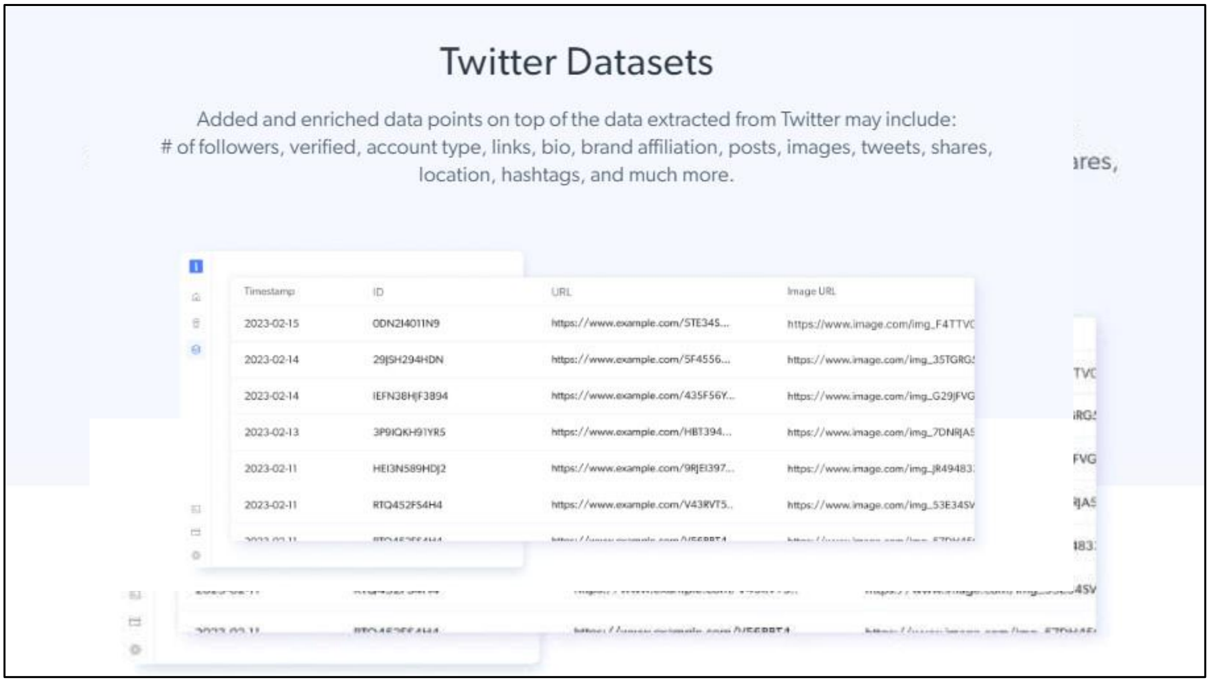
Request dataset >

Username	Following	Followers	Posts	Shares	Verified account
@MarkWilliams	23K	36K	4K	85K	Yes
@JonathanSmith	97K	120K	13K	2M	Yes
@AmandaMiller	49K	60K	7K	92K	No

~~*See* Exh. E.~~

~~47.~~116. According to Defendant's website, the X Corp. data sets offered for sale by Defendant include "millions of pages and tens of millions of data points." Specifically, these data sets include the following user information: "# of followers, verified, account type, links, bio, brand affiliation, posts, images, tweets, shares, location, hashtags, and much more."

Figure ~~8~~9: Screenshot of Bright Data's website on July 10, 2023



Twitter Datasets

Added and enriched data points on top of the data extracted from Twitter may include:
# of followers, verified, account type, links, bio, brand affiliation, posts, images, tweets, shares, location, hashtags, and much more.

Timestamp	ID	URL	Image URL
2023-02-15	ODN2I4011N9	https://www.example.com/STE345...	https://www.image.com/img_F4TTVC
2023-02-14	29JSH294HDN	https://www.example.com/5F4556...	https://www.image.com/img_35TGRG
2023-02-14	IEFN38HJF3894	https://www.example.com/435F56Y...	https://www.image.com/img_G29JFVG
2023-02-13	3P9IQKH91YR5	https://www.example.com/HBT394...	https://www.image.com/img_7DNRJAS
2023-02-11	HEI3N589HDJ2	https://www.example.com/9RJEI397...	https://www.image.com/img_JR49483
2023-02-11	RTQ452F54H4	https://www.example.com/V43RVT5...	https://www.image.com/img_53E34SV

~~*See id.*~~

~~48.~~117. Defendant could have only obtained this data by engaging in prohibited scraping of X's platform.

Defendant offers this unlawfully obtained data for sale starting at $.01 per record, but also offers customized packages of X Corp.'s data.

~~49.~~118. Defendant also offers several options for delivery of X Corp.'s data, and even offers its customers the opportunity to ~~regularly~~ update its data sets with additional data scraped from X at regular intervals.

119. Defendant does not transform the content it scrapes from X, nor does it offer that content for social or educational purposes. Rather, as Defendant's website makes clear, its purpose is commercial: it packages up X content and data and sells it to the highest bidder. Indeed, the "Twitter Datasets" Bright Data sells merely duplicate the data that X's own API tiers already make available, without X's privacy and other safeguards. In doing so, Bright Data does not obtain consent from X's users before scraping and selling their data for commercial purposes. On information and belief, X's users broadly remain unaware that Bright Data is doing so.

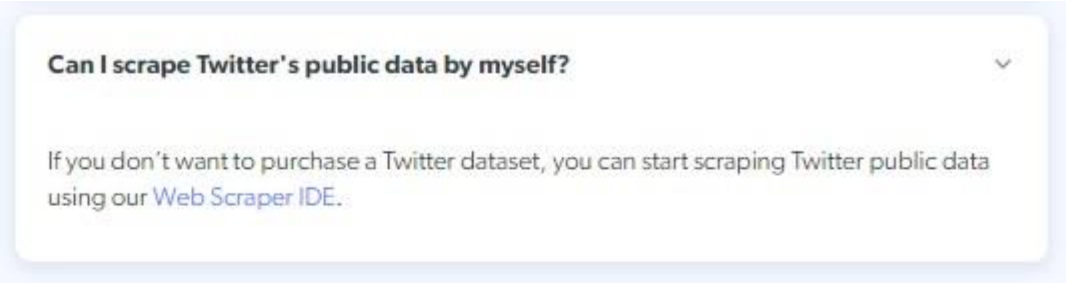
120. Defendant claims dubiously that its datasets sourced from X contain only publicly accessible data – that is, data available to users not logged into X. On information and belief, that claim is inaccurate for the reasons stated above. But even if it were true, there is no way Bright Data could have obtained this amount of data without circumventing the technological measures X put in place to prevent scraping of data available to users that are not logged into X. Specifically, Bright Data would have needed at the very least to circumvent X's IP-based rate limits through proxy networks, and to circumvent X's anomaly detection measures. Without doing so, Bright Data could not have obtained the high volume of data it sells, because of the stringent limits X places on the amount of data that can be accessed before being required to login or create an account.

2. Bright Data Sells Automated Tools to Scrape X Corp.'s Data

~~50.~~121. Defendant also offers for sale on its website automation software that allows users to scrape data directly from the X platform in violation of X Corp.'s Terms.

~~51.~~122. ~~As indicated in Exhibit F,~~Defendant is aware that its customers use its tools to scrape data from X through logged-in accounts. On information and belief, Bright Data tools have the capability to determine whether its tools are being used to login to X's platform and could stop such use if Bright Data wanted. However, Bright Data knowingly allows logged-in scraping to occur so long as Bright Data users comply with a sham "Know Your Customer" process. Defendant's website states: "If you don't want to purchase a Twitter dataset, you can start scraping Twitter ~~public~~ data using our ~~Web Scraper IDE."~~Twitter scraping tool." Bright Data, *Twitter Datasets*, https://perma.cc/L72J-WGVL.

~~**Figure 9: Screenshot of Bright Data's website on July 10, 2023**~~



Can I scrape Twitter's public data by myself?

If you don't want to purchase a Twitter dataset, you can start scraping Twitter public data using our Web Scraper IDE.

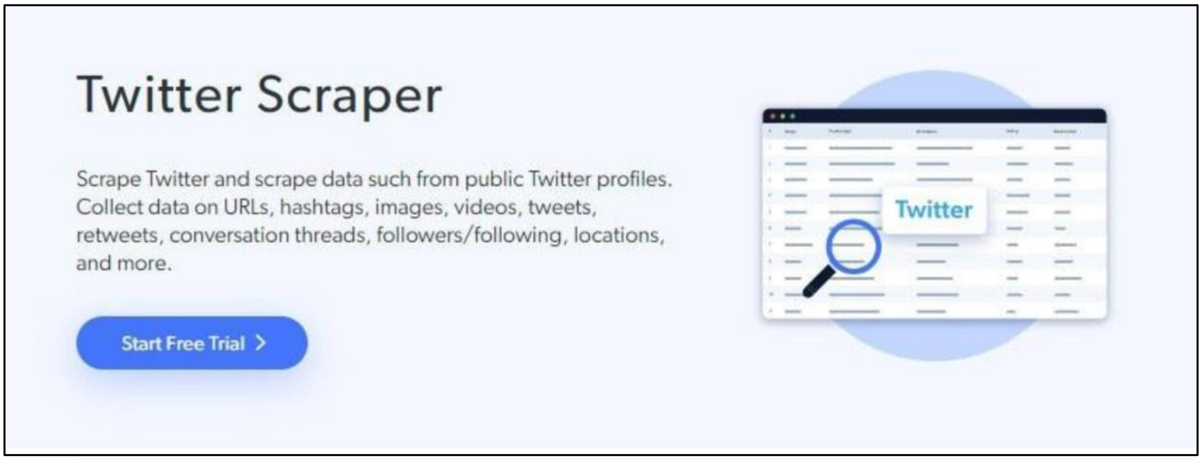
~~*See* Exh. E.~~

~~52.~~123. Defendant's Web Scraper tool allows individuals to evade detection utilizing a proxy network in order "to remain anonymous, avoid IP blocking, access geo-restricted content, and improve scraping speed." Bright Data, *Twitter Scraper*, https://perma.cc/4EFH-PKVT. The tool also includes an "unblocking solution" that is designed to evade anti-scraping measures like those employed by X Corp. *Id.* Defendant specifically advertises that its Web Scraper tool can be used to "[e]asily scrape data from any geo-location while avoiding CAPTCHAs and blocks." Bright Data, *Web Scraper APIs*, https://perma.cc/MHA9-Q3GG.

~~53.~~124. In addition to its Web Scraper tool, Defendant sells at least four additional tools designed to scrape information specifically from the X Platform: Twitter Scraper, Twitter Profile Scraper, Twitter Image Scraper, and Twitter Followers Scraper.

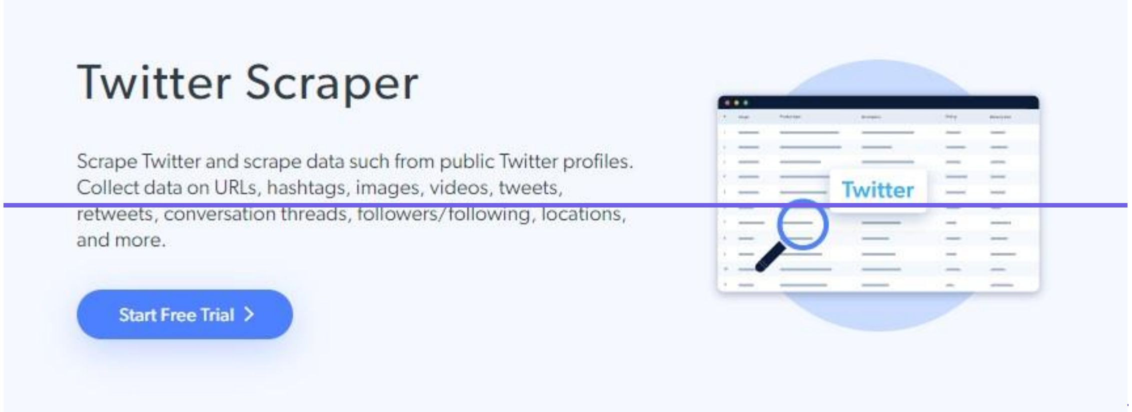
a. ~~As seen in Figure 7 below,~~ Defendant offers a Twitter Scraper to automatically scrape data from the X platform, including "URLs, hashtags, images, videos, tweets, retweets, conversation threads, followers/following, locations, and more." Bright Data, *Twitter Scraper API*, https://perma.cc/49MC-3HUL. Bright Data advertises its Twitter Scraper's ability to circumvent X's defenses: "Maintain full control, flexibility, and scale without worrying about infrastructure, proxy servers, or getting blocked." *Id.*

Figure 10: Screenshot of Bright Data's ~~website~~Website on July 10, 2023



Twitter Scraper

Scrape Twitter and scrape data such from public Twitter profiles. Collect data on URLs, hashtags, images, videos, tweets, retweets, conversation threads, followers/following, locations, and more.



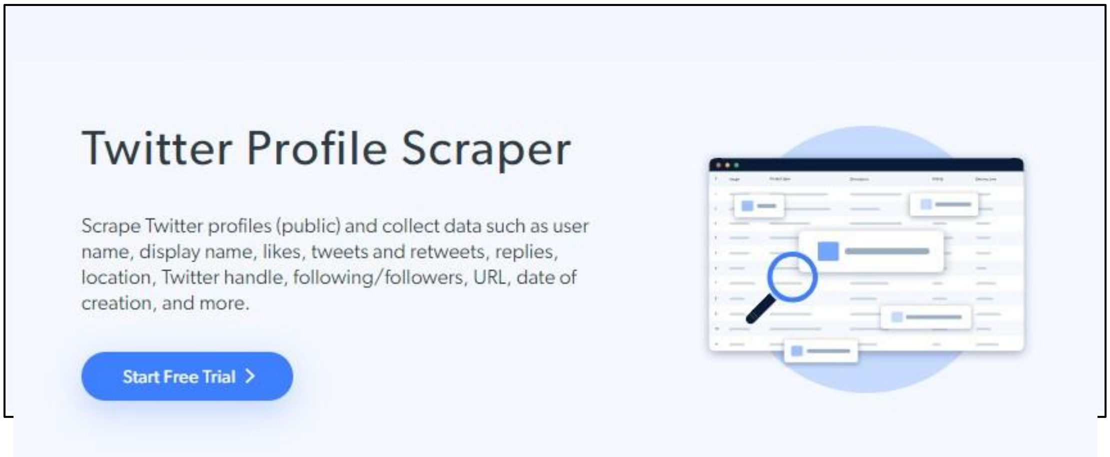
~~Twitter Scraper~~

~~Scrape Twitter and scrape data such from public Twitter profiles. Collect data on URLs, hashtags, images, videos, tweets, retweets, conversation threads, followers/following, locations, and more.~~

~~*See* Exh. F.~~

b. ~~As seen in Figure 11 below,~~ Defendant offers a Twitter Profile Scraper to automatically "collect data such as user name, display name, likes, tweets and retweets, replies, location, Twitter handle, following/followers, URL, date of creation, and more." *See also* Bright Data, *Twitter Profile Scraper API*, https://perma.cc/3TAQ-VGVC.

Figure 11: Screenshot of Bright Data's website on July 10, 2023



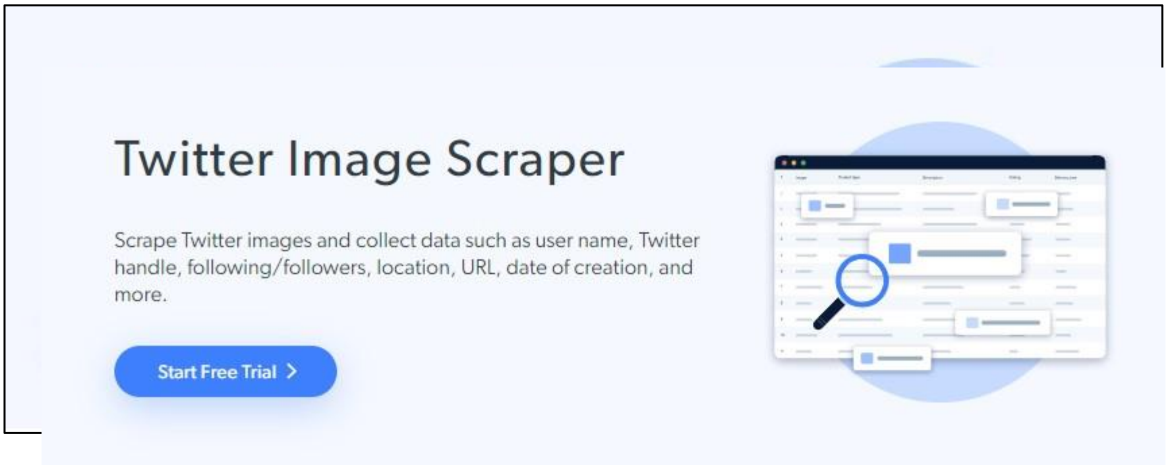
Twitter Profile Scraper

Scrape Twitter profiles (public) and collect data such as user name, display name, likes, tweets and retweets, replies, location, Twitter handle, following/followers, URL, date of creation, and more.

~~*See* Exh. G.~~

c. ~~As seen in Figure 12 below,~~ Defendant also offers a Twitter Image Scraper to automatically "collect data such as user name, Twitter handle, following/followers, location, URL, date of creation, and more." *See also* Bright Data, *Twitter Images Scraper API*, https://perma.cc/L8SX-W5WC.

Figure 12: Screenshot of Bright Data's website on July 10, 2023



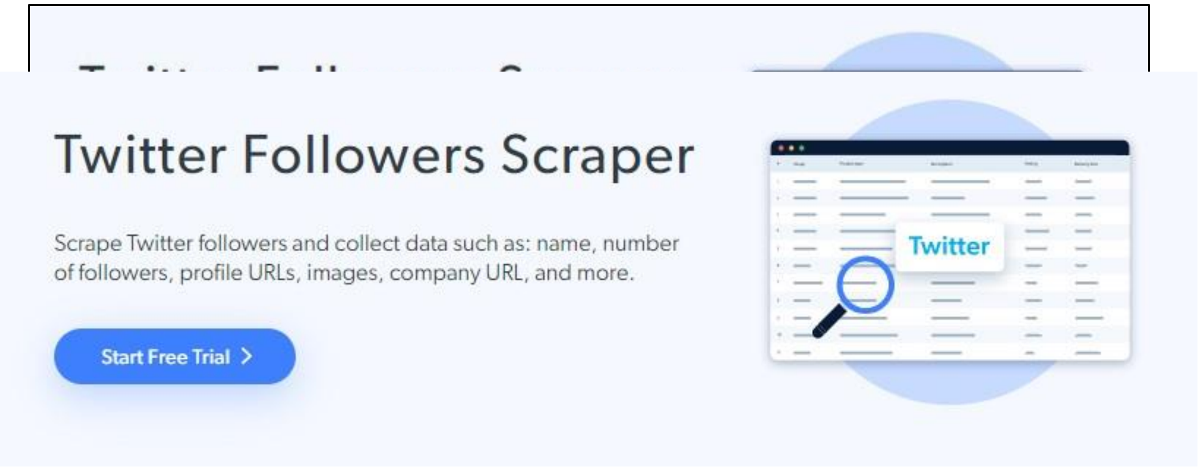
Twitter Image Scraper

Scrape Twitter images and collect data such as user name, Twitter handle, following/followers, location, URL, date of creation, and more.

~~*See* Exh. H.~~

d. ~~As seen in Figure 13 below,~~ Defendant has also offered a Twitter Followers Scraper to automatically collect data such as "name, number of followers, profile URLs, images, company URL, and more." *See also* Bright Data, *Twitter Followers Scraper*, https://perma.cc/92LX-4PVK.

Figure 13: Screenshot of Bright Data's website on July 10, 2023



Twitter Followers Scraper

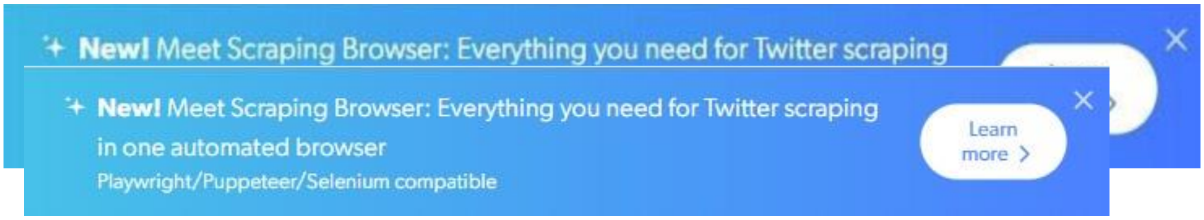
Scrape Twitter followers and collect data such as: name, number of followers, profile URLs, images, company URL, and more.

~~*See* Exh. I.~~

~~54.~~125. For each of these products, Defendant claims it "[u]tilizes proprietary technology to unlock sites" and allows customers to "collect as much data as you need quickly and completely."

~~55.~~126. In addition to these X-specific scraping tools, Bright Data offers an automated "Scraping Browser" that simplifies the act of scraping data from the X platform. ~~As seen in Figure 14 below,~~ Bright Data markets this product for scraping X Corp.'s data.

Figure 14: Screenshot of Bright Data's website on July 10, 2023



+ New! Meet Scraping Browser: Everything you need for Twitter scraping

+ New! Meet Scraping Browser: Everything you need for Twitter scraping in one automated browser

Playwright/Puppeteer/Selenium compatible

~~*See* Exh. G.~~

~~56.~~127. Defendant advertises this "Scraping Browser" as containing "all website unlocking operations under the hood, including: CAPTCHA solving, browser fingerprinting, automatic retries, selecting headers, cookies, & Javascript rendering, and more." Defendant also claims its Scraping Browser "automatically learns to bypass bot-detection systems as they adapt, saving you the hassle and cost."

~~57.~~128. The Scraping Browser allows Defendant's customers to "appear as a real user browser to bot-detection system[s]," such as those used by X Corp.

3. Bright Data Sells Proxy Services to Facilitate Data-Scraping

129. Bright Data touts that it has more than 72 million "residential IPs" available for use on a rotating basis to support "unlimited concurrent sessions" by data scrapers – that is, massive coordinated scraping. Bright Data, *Residential Proxies*, https://perma.cc/PM7C-75UZ. Bright Data appears to "source" those IPs in two ways. First, it persuades individuals to download EarnApp by paying them incentives; the app then allows Bright Data to route data scrapers' requests through those individuals' IPs to hide the true source of the scraping requests to X's servers. Second, Bright Data bundles its "Bright SDK" with other apps, and when individuals install those other apps, Bright Data can again route data scrapers' requests through those individuals' IPs to hide the true source of the traffic.

~~58.~~130. Defendant also facilitates ~~the violation~~violations of X Corp.'s Terms by offering proxy services specifically designed to evade anti-scraping measures, including X Corp.'s CAPTCHAs and its user ID and IP rate limits. These tools allow unregistered users to impersonate ~~actual~~registered X users ~~in order~~ to bypass X Corp.'s ~~defenses~~digital fence and gates.

~~59.~~131. These proxy services imitate requests from legitimate users ~~in order~~ to conceal the true requestor's IP address and location. Defendant advertises that these proxies will "avoid[] IP bans and CAPTCHAs" and allow users to "[g]ather vast amounts of public web data with total anonymity." ~~*See* Exh. J.~~

132. Bright Data's Proxy Services are intentionally deceptive. Bright Data does not allow data scrapers merely to use one proxy IP – like, for example, a user in another country simulating a U.S. person to access content available in the United States. Rather, Bright Data provides millions of proxies that can be used concurrently and on a rapidly rotating basis. These rotating proxies allow a single data scraper to spread millions of requests across many different IPs and fake accounts. The deceptive proxy services also impede X's efforts to ban requests by individual IPs because, as soon as one IP is banned, Bright Data's software automatically routes the request through a series of new IPs. Although Bright Data's proxy servers make this nearly impossible to track, on information and belief X has blocked numerous accounts for scraping and other violations of its Terms that Bright Data's proxy servers then allowed to regain access to the platform through a different IP. *See* Bright Data, *Proxy Solutions*, https://perma.cc/CL2J-3L3T

(“When using Bright Data’s award-winning rotating proxies, you can scrape data from any website in the world with a 99.99% success rate. These rotating proxies will constantly replace your IP address, ensuring that you won’t get flagged or blocked.”).

133. As Bright Data explains on its website, a “rotating proxy is a proxy server that rotates your IP each time you connect through the proxy server, using proxies from a massive pool of millions of IPs.” *See* Bright Data, *Rotating Proxies*, https://perma.cc/S264-G97X. This allows “limitless concurrent connections” that can be used to scrape massive amounts of data. *See* Bright Data, *Residential Proxies*, https://perma.cc/PM7C-75UZ.

134. There is no legitimate purpose for such tools other than to circumvent X’s technological measures. Ordinary proxy services, like a VPN, might be used to maintain a user’s privacy in countries with oppressive regimes. That is not what Bright Data markets and sells.

135. The identity masking that Bright Data’s proxy services facilitate further harm X and its users by impeding X from fulfilling its obligations to give users control over their data, including under statutes like General Data Protection Regulation and the California Consumer Privacy Act that require companies to delete consumer data upon request. X cannot ensure that its users’ data is deleted once it has been anonymously scraped. By contrast, if the data is obtained through X’s API, X has a business relationship with the developer and can demand the deletion of data through established contractual processes.

136. Bright Data does not in its Acceptable Use Policy provide analogous privacy features and control over X user data scraped from the X platform:

a. Bright Data does not require itself or its customers to delete or make private any data scraped from X after an X user has deleted or privatized it;

b. Bright Data does not prohibit tracking of individuals based on the characteristics which X protects;

c. Bright Data does not prohibit matching X usernames to users’ legal identities or personally identifying information;

d. Bright Data does not prohibit users from using geodata to track X users, including the creation of heat maps or user location profiles, even in circumstances in which that information is potentially highly sensitive;

e. Bright Data does not restrict scraping from high-risk jurisdictions or take any steps to ensure that X's data is not used by malign actors;

f. Bright Data does not prohibit its scraped data from being used (including by foreign governments) to assist in election interference, voter suppression, or the tracking and targeting of sensitive groups, including activists and political dissidents; and

g. Bright Data does not prohibit its scraped data being used for individual profiling, psychographic segmentation, background and credit checks, or the development of facial recognition.

FIRST CAUSE OF ACTION

(Breach of Contract)

~~60.~~137. X Corp. realleges and incorporates all preceding paragraphs herein.

~~61.~~138. Use of the X platform and use of X Corp.'s services are governed by X Corp.'s Terms.

~~62.~~139. X users, including Defendant, accept the Terms as a condition of using the platform.

~~63.~~140. Moreover, by virtue of having X accounts, Defendant has expressly accepted and agreed to X Corp.'s Terms.

~~64.~~141. The Terms are enforceable and binding on Defendant.

~~65.~~142. Defendant has repeatedly violated the Terms, including by (i) accessing the X platform through automated means without ~~specific~~ authorization from X Corp.; (ii) scraping data from the X platform without authorization; (iii) selling tools that enable others, including X users, to access the X platform by automated means and to scrape data; (iv) selling proxy services that enable others, including X users, to access the X platform by automated means and evade X Corp.'s anti-automation and anti-scraping tools; and (v) selling data that Defendant scraped from the X platform.

~~66.~~143. Defendant has breached and continues to breach the Terms by scraping data from X Corp.'s platform without prior consent from X Corp. X Corp. has never authorized

Defendant to access its platform through automated means and has never given Defendant consent to scrape data.

~~67.~~144. Despite being bound by the Terms, Defendant has repeatedly accessed the X Corp. platform through automated means and scraped data in violation of the Terms.

~~68.~~145. Defendant has breached, and continues to breach, X Corp.'s Terms by accessing the platform through unauthorized means and scraping data from the platform.

~~69.~~146. Defendant has breached, and continues to breach, X Corp.'s Terms by selling tools that allow other X users to access the platform by automated means and scrape data, and by selling proxy services that allow the same.

~~70.~~147. Defendant has breached, and continues to breach, X Corp.'s Terms by selling data that Defendant has scraped from X Corp.'s platform.

~~71.~~148. Defendant's conduct – both accessing X Corp.'s platform in volumes and manners that violate the Terms as well as selling data scraped from X Corp.'s platform – has damaged X Corp. and caused and continues to cause irreparable harm and injury to X Corp.

~~72.~~149. X Corp. is entitled to compensatory damages, injunctive relief, declaratory relief, and/or other equitable relief.

SECOND CAUSE OF ACTION

(Tortious Interference with Contract)

~~73.~~150. X Corp. realleges and incorporates all preceding paragraphs herein.

~~74.~~151. All X users must agree to abide by the Terms, which constitute a valid and enforceable agreement between X Corp. and each user.

~~75.~~152. As a user of X Corp.'s platform, as well as a present or former X account holder, Defendant is aware of the Terms and that they govern all users who choose to interact with the X platform. Defendant is also aware of the Terms because several of its executives and employees are present or former X account holders.

~~76.~~153. Nevertheless, Defendant has marketed and sold its scraping tools to X users and account holders, including X users and account holders residing in California, through its interactive website accessible in California and elsewhere, through its sales office and employees in California and elsewhere, and by using the X platform to market its scraping services to other X users and account holders.

~~77.~~154. Defendant has also sold proxy services and tools to facilitate automated access and scraping of the X platform by X users and account holders, including by locally offering a "Superior California Proxy" with "[v]ast numbers of California IPs to get data off any website."

~~78.~~155. By offering services and tools designed to provide automated access to the X platform, and to scrape data from the platform, Defendant induced a breach or disruption of the Terms by X users.

~~79.~~156. On information and belief, those who purchased Defendant's scraping services and tools used them to access X through unauthorized, automated means and to scrape data from the X platform, in violation of the Terms.

~~80.~~157. Defendant's conduct has damaged X Corp. and caused and continues to cause irreparable harm and injury to X Corp.

~~81.~~158. X Corp. is entitled to compensatory damages, injunctive relief, declaratory relief, and/or other equitable relief.

THIRD CAUSE OF ACTION

(Unjust Enrichment, in the alternative)

~~82.~~159. X Corp. realleges and incorporates all preceding paragraphs herein.

~~83.~~160. If Defendant's acts are found not to be in breach of contract, then Defendant's acts as alleged herein constitute unjust enrichment at X Corp.'s expense.

~~84.~~161. Defendant used X Corp.'s service, platform, and computer network without authorization to scrape data from the X platform.

~~85.~~162. Defendant receives benefits in the form of profits from its unauthorized scraping of ~~X Corp.~~ data~~.~~ from the X platform.

~~86.~~163. Defendant's retention of the profits derived from its unauthorized scraping of data would be unjust.

~~87.~~164. Defendants' conduct has damaged X Corp., including but not limited to hampering the user experience for authentic X users and customers, in addition to the time and money spent investigating and mitigating Defendants' unlawful conduct.

~~88.~~165. X Corp. seeks actual damages from Defendants' unlawful activities, an accounting, and disgorgement of Defendants' profits in an amount to be determined at trial, compensatory damages, injunctive relief, declaratory relief, and/or other equitable relief.

FOURTH CAUSE OF ACTION

(Trespass to Chattels)

~~89.~~166. X realleges and incorporates all preceding paragraphs herein.

~~90.~~167. The X platform and all underlying technological infrastructure are the personal property of X Corp.

~~91.~~168. Defendant intentionally entered into, and made use of, X Corp.'s technological infrastructure, including its software and servers located in California, to obtain information for its own economic benefit.

~~92.~~169. Defendant knowingly exceeded the permission granted by X Corp. to access its personal property, including its technological infrastructure and servers.

~~93.~~170. Defendant's acts have diminished the server capacity that X Corp. can devote to its legitimate users, ~~and~~ thereby ~~injured~~injuring X Corp. by ~~depriving it~~requiring X Corp. to purchase additional service capacity and diminishing the condition and quality of ~~the ability~~X Corp.'s service to ~~use~~ its ~~personal property.~~ legitimate users.

~~94.~~171. Through its acts, Defendant also caused other persons, including X users and account holders based in California and elsewhere, to knowingly exceed the permission granted by X Corp. to access its personal property, further injuring X Corp.

~~95.~~172. X Corp. has never consented to Defendant's conduct.

~~96.~~173. Defendant's conduct constitutes trespass to X Corp.'s chattels.

~~97.~~174. Defendant's acts have caused injury to X Corp. and if continued, expanded, and/or replicated unchecked by others, will cause damage in the form of impaired condition, quality, and value of its servers, technology infrastructure, services, and reputation.

FIFTH CAUSE OF ACTION

(Unlawful, Unfair, or Fraudulent Business Practices (Cal. Bus. & Prof. Code § 17200 et seq.))

~~98.~~175. X Corp. realleges and incorporates all preceding paragraphs herein.

176. Defendant's actions described above constitute unlawful, unfair, ~~or~~and fraudulent acts or practices in the conduct of a business, in violation of California's Business and Professions Code Section 17200, et seq~~, including~~.

~~99.~~177. Defendant's actions violate the Unfair Competition Law's ("UCL") "unlawful" prong because they constitute trespass and tortious interference with business relationships in violation of the law, and ~~because Defendant deceived X Corp. into providing it access to, and information from, the X Corp. computer network. Defendant's data-collection technology and its data-scraping tools deliberately misrepresented the requests sent to the X platform, posing as legitimate X users, and Defendant's sale of IP proxies masquerades as a legitimate X user to avoid X Corp.'s technical measures designed to prevent unauthorized access of its computer servers~~violations of the DMCA, CFAA, and CDAFA.

178. Defendant's actions violate the UCL's "unfair" prong because they are unethical, oppressive, unscrupulous, or substantially injurious to consumers and offend established public policy or are immoral. For instance, Defendant circumvents anti-scraping measures designed to protect consumers' privacy rights, including those under the California Consumer Privacy Act, the European Union's General Data Protection Regulation, and other acts with which X Corp. expends substantial resources to comply. The availability of unrestricted third-party access to scraped data also harms X's ability to accurately and effectively prevent and deter market manipulation, scams, and fraud.

179. Defendant's actions violate the UCL's "fraudulent" prong because Defendant and its customers have engaged in widespread scraping of data that is accessible only to X users, developers, or advertisers who are logged into registered, password-protected accounts, and have evaded X Corp.'s blocking measures by, among other things, creating new fake accounts to engage in the same scraping activity for which a prior account was blocked.

~~100.~~180. Scraping data, as well as circumventing X Corp.'s ability to police its own platform, has caused substantial injury to X Corp., in the form of costs to investigate, remediate, and prevent Defendant's wrongful conduct, among other injuries.

~~101.~~181. As a result of Defendant's various acts and omissions, X Corp. has suffered and continues to suffer irreparable harm for which there is no adequate remedy at law, and which will continue unless Defendant's actions are enjoined.

SIXTH CAUSE OF ACTION

(Misappropriation)

~~102.~~182. X Corp. realleges and incorporates all preceding paragraphs herein.

~~103.~~183. X Corp. has invested substantial time, labor, skill, and financial resources into the creation and maintenance of X, its computer systems, and servers, including system and server capacity, as well as the ~~content on X.~~aggregated data at scale. Defendant has not invested any of its own time nor resources to the development of the X platform.

~~104.~~184. Defendant used automated means~~—~~ – in violation of X Corp.'s Terms~~—~~ – to wrongfully access the X platform, systems and servers, including systems and servers located in California, and obtain aggregated data at scale from the X platform.

~~105.~~185. Defendant appropriated this aggregated data at scale at little or no cost to Defendant, free-riding on X Corp.'s substantial investment of time, effort, and expense to aggregate this data at scale.

~~106.~~186. As a result of Defendant's misappropriation, X Corp. has been forced to expend additional time, labor, skill and financial resources to investigate and mitigate Defendant's wrongful conduct. Defendant has been able to exploit and profit from X Corp.'s substantial investments in the X platform and the creation of its aggregated data at scale.

~~107.~~187. X Corp. has been and will continue to be damaged as a result of Defendant's misappropriation.

~~108.~~188. X Corp. has suffered and will continue to suffer irreparable injury, and its remedy at law is not itself adequate to compensate it for injuries inflicted by Defendant.

SEVENTH CAUSE OF ACTION

(Violation of the DMCA, 17 U.S.C. § 1201(a)(1)(A) and (a)(2))

189. X Corp. realleges and incorporates all preceding paragraphs herein.

190. X Corp. owns valid copyrights in its websites, including twitter.com and X.com, and mobile and online applications.

191. X Corp. employs technological measures to control access to its websites and applications. These measures include CAPTCHA, login requirements, rate limits, robots.txt restrictions, and anomaly detection tools.

192. These measures require the application of information, or a process or a treatment to gain access to X Corp.'s websites and applications. They require inputting information (CAPTCHA), providing a valid username and password (logins), identifying the IP and/or account of the individual requesting access (rate limits), and use of the service that reflects human use as opposed to automated access (anomaly detection tools).

193. Defendants' tools – including Proxy Solutions, Scraping Browser, Web Unlocker, and Scraper API – circumvent these technological measures by avoiding, bypassing, and impairing each of these measures. That is the primary design and only commercially significant purpose and use of these tools. Defendants' tools automatically crack CAPTCHA prompts; they enable mass use of accounts to bypass rate limits on any individual account; they provide millions of IP addresses on a rotating basis to evade IP-based rate limits; and they mimic human behavior through automated means to avoid anomaly detection, too.

194. Defendant has used and continues to use automated systems to engage in widespread scraping of data on the X platform, repeatedly bypassing, avoiding, disabling, deactivating, or impairing the technological measures controlling access to X Corp.'s copyrighted websites and applications in violation of 17 U.S.C. § 1201(a)(1)(A).

195. Defendant also offers its tools to the public. These tools are primarily designed to circumvent X Corp.'s technological measures, and do not have more than limited commercially significant purposes other than circumventing X Corp.'s technological measures. Defendant openly markets these tools as being available to circumvent technological measures, including X Corp.'s specifically. This violates 17 U.S.C. § 1201(a)(2)(A).

196. Defendant's acts constituting DMCA violations have been and continue to be performed without the authorization or consent of X Corp. and in violation of its Terms.

197. Defendant has violated Section 1201 of the DMCA willfully.

198. Defendant's conduct has caused damage to X Corp., in the form of costs to investigate, remediate, and prevent Defendant's wrongful conduct, among other injuries, and has unjustly enriched Defendant.

199. X Corp. has suffered and will continue to suffer irreparable injury, and its remedy at law is not itself adequate to compensate it for injuries inflicted by Defendant.

EIGHTH CAUSE OF ACTION

(Violation of the Computer Fraud & Abuse Act (CFAA)

18 U.S.C. § 1030(a)(2)(C) and (a)(4))

200. X Corp. realleges and incorporates all preceding paragraphs herein.

201. The CFAA provides that "[w]hoever . . . intentionally accesses a computer without authorization or exceeds authorized access, and thereby obtains . . . information from any protected computer," is subject both to criminal and civil liability. 18 U.S.C. § 1030(a)(2)(C).

202. Defendant has repeatedly and intentionally accessed X Corp.'s servers without authorization and in violation of X Corp.'s Terms and has continued to do so even after X Corp. filed this lawsuit.

203. X Corp.'s servers are a "computer" within the meaning of the CFAA, which defines that term to include "any data storage facility or communications facility directly related to or operating in conjunction with [a computer]." *Id.* § 1030(e)(1).

204. X Corp.'s servers also constitute a "protected computer" within the meaning of the CFAA, because they are connected to the Internet and are used in and affect interstate and foreign commerce and communications. *Id.* § 1030(e)(2)(B).

205. Defendant has circumvented X Corp.'s technological barriers and access restrictions, including CAPTCHAs, login requirements, rate limits, robots.txt restrictions, and anomaly detection tools to engage in widespread scraping of non-public data on X Corp.'s servers that is accessible only to X users, developers, or advertisers who are logged into registered, password-protected accounts.

206. Defendant has also sold and advertised tools – including Proxy Solutions, Scraping Browser, Web Unlocker, and Scraper API – that circumvent X Corp.'s technological

barriers and access restrictions and facilitate the scraping of non-public data on the X platform, thereby directing, encouraging, or inducing others to intentionally access X Corp.'s servers without authorization in violation of the CFAA.

207. The CFAA also prohibits "knowingly and with intent to defraud, access[ing] a protected computer without authorization, or exceed[ing] authorized access, and by means of such conduct further[ing] the intended fraud and obtain[ing] anything of value." *Id.* § 1030(a)(4).

208. By engaging in misrepresentations, Defendant directly or indirectly accesses X Corp.'s platform and computer network to evade anti-scraping measures and scrape data that is accessible only to X users, developers, or advertisers who are logged into registered, password-protected accounts.

209. Those misrepresentations include using "straw man" accounts and millions of rotating and deceptive IP addresses that mask the true requester. For instance, Defendant's tools circumvent X Corp.'s rate limits and usage restrictions by flooding X Corp.'s servers with requests all originating from the same entity but routed through millions of different IP addresses and/or many different user accounts to conceal that the same entity is making the request.

210. Defendant has caused X Corp. substantial damages and losses, including, without limitation, harm caused by the increased burden on and interruption of service to X Corp.'s website, data and/or underlying databases, amounts expended attempting to prevent the Defendant's unauthorized scraping, and other losses in an amount well over $5,000 aggregated over a one-year period.

211. X Corp. has suffered and will continue to suffer irreparable injury, and its remedy at law is not itself adequate to compensate it for injuries inflicted by Defendant.

NINTH CAUSE OF ACTION

(Violations of the California Comprehensive Computer Data Access and Fraud Act (CDAFA) California Penal Code § 502)

212. The CDAFA provides for criminal and civil liability against "any person" who, among other specified misconduct, (a) "[k]nowingly accesses and without permission takes,

copies, or makes use of any data from a computer, computer system, or computer network, or takes or copies any supporting documentation, whether existing or residing internal or external to a computer, computer system, or computer network"; (b) "[k]nowingly and without permission uses or causes to be used computer services"; (c) "[k]nowingly and without permission provides or assists in providing a means of accessing a computer, computer system, or computer network in violation of this section"; or (d) "[k]nowingly and without permission accesses or causes to be accessed any computer, computer system, or computer network." Cal. Penal Code § 502(c).

213. X Corp.'s servers constitute and are made up of one or more "computer networks," "computer systems," and "computer programs" or "software," and provide "computer services" to X users, developers, and advertisers.

214. A substantial portion of these networks and systems is located in the State of California.

215. As set forth above, Defendant has repeatedly and knowingly accessed X Corp.'s servers without X Corp.'s permission and in violation of its Terms and anti-scraping measures, including in California. Among other things, Defendant has circumvented, and caused others to circumvent, X Corp.'s technical and code-based barriers that restrict scraping of non-public data on X's platform.

216. Defendant has caused X Corp. substantial damages and losses, including, without limitation, harm caused by the increased burden on and interruption of service to X Corp.'s website, data and/or underlying databases, amounts expended attempting to prevent the Defendant's unauthorized scraping, and other losses and damage.

217. X Corp. has suffered and will continue to suffer irreparable injury, and its remedy at law is not itself adequate to compensate it for injuries inflicted by Defendant.

PRAYER FOR RELIEF

WHEREFORE, Plaintiff prays for relief, as follows:

1. Preliminary and permanent injunctive relief enjoining Bright Data, its agents, officers, employees, and successors from:

a. accessing or using X Corp.'s website, servers, systems, and any data contained therein for purposes of unlawful data scaping;

b. developing or distributing any technology or product that is used, or could be used, for the unauthorized scraping of data from X;

c. facilitating the scraping of data by other users;

d. selling or offering for sale any data previously obtained from X;

e. utilizing any proxies to access X's website, servers, systems, and any data contained therein; and

f. selling or offering for sale any proxies that can be used to access X's website, servers, systems, and any data contained therein.

2. That Defendant be required to identify the location of any and all data obtained from the X platform and to destroy any and all such data;

3. That Defendant be required to identify any and all recipients of data obtained from the X platform;

4. Compensatory, statutory, and punitive damages, as permitted by law and in such amounts to be proven at trial;

5. Reasonable costs, including reasonable attorneys' fees;

6. Pre- and post-judgment interest, as permitted by law;

7. An accounting of Defendant's profits from its scraping activities and disgorgement ~~of~~ those profits; and

8. Any other remedy to which Plaintiff X Corp., Inc. may be justly entitled.

~~Dated: November 14, 2023~~ ~~Respectfully submitted,~~

~~**HAYNES & BOONE LLP**~~

~~By: /s/ Jason T. Lao~~

~~David H. Harper*~~
~~david.harper@haynesboone.com~~
~~Jason P. Bloom*~~
~~jason.bloom@haynesboone.com~~

~~2801 N. Harwood Street~~
~~Suite 2300~~
~~Dallas, Texas 75201~~
~~Telephone: (214) 651.5000~~
~~Telecopier: (214) 651.5940~~
~~* *Admitted Pro Hac Vice*~~

~~Jason T. Lao~~
~~jason.lao@haynesboone.com~~
~~Andrea Levenson~~
~~andrea.levenson@haynesboone.com~~
~~600 Anton Boulevard, Suite 700~~
~~Costa Mesa, California 92626~~
~~Telephone: (949) 202-3000~~
~~Facsimile: (949) 202-3001~~

~~*Attorneys for Plaintiff X Corp.*~~

DATED: August 16, 2024

Respectfully submitted,

KELLOGG, HANSEN, TODD,
FIGEL & FREDERICK, P.L.L.C.

By: /s/ Joshua D. Branson
JOSHUA D. BRANSON
DANIEL V. DORRIS
BETHAN R. JONES
MATTHEW D. READE
TIBERIUS T. DAVIS

Attorneys for Plaintiff
X Corp.

DEMAND FOR JURY TRIAL

Pursuant to Federal Rule of Civil Procedure 38(b), Plaintiff hereby demands a trial by jury of all triable issues.

~~Dated: November 14, 2023~~ ~~**HAYNES AND BOONE LLP**~~

~~By: /s/ Jason T. Lao~~

~~Jason T. Lao~~

DATED: August 16, 2024

Respectfully submitted,

KELLOGG, HANSEN, TODD,
FIGEL & FREDERICK, P.L.L.C.

By: /s/ Joshua D. Branson
JOSHUA D. BRANSON
DANIEL V. DORRIS
BETHAN R. JONES
MATTHEW D. READE
TIBERIUS T. DAVIS

Attorneys for Plaintiff
X Corp.